



Techniques de corrélation : aspects méthodologiques et applications à la sub-surface

Pierre Gouedard

► To cite this version:

Pierre Gouedard. Techniques de corrélation : aspects méthodologiques et applications à la sub-surface. Géophysique [physics.geo-ph]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT : . tel-00357045

HAL Id: tel-00357045

<https://theses.hal.science/tel-00357045>

Submitted on 29 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ JOSEPH FOURIER – GRENOBLE I

OBSERVATOIRE DES SCIENCES DE L'UNIVERS DE GRENOBLE

LABORATOIRE DE GÉOPHYSIQUE INTERNE ET DE TECTONOPHYSIQUE

TECHNIQUES DE CORRÉLATION : ASPECTS MÉTHODOLOGIQUES ET APPLICATIONS À LA SUB-SURFACE



Thèse présentée par

Pierre GOUÉDARD

le 8 décembre 2008

en vue de l'obtention du titre de

Docteur de l'Université Joseph Fourier – Grenoble I

École doctorale Terre, Univers, Environnement

Composition du jury :

Philippe ROUX	Ch. de Recherche au LGIT, Grenoble	Directeur de thèse
Michel CAMPILLO	Pr. à l'Univ. J. Fourier, LGIT, Grenoble	Directeur de thèse
Philippe CÔTE	Dir. de Recherche au LCPC, Nantes	Rapporteur
Nikolai SHAPIRO	Dir. de Recherche à l'IPG, Paris	Rapporteur
Anne PAUL	Dir. de Recherche au LGIT, Grenoble	Examinatrice
Arie VERDEL	Research Geophysicist at Shell, The Hague	Examineur
Richard WEAVER	Pr. at Illinois University, Urbana	Examineur

RÉSUMÉ

Les méthodes traditionnelles d'analyse du sous-sol reposent sur l'utilisation de sources sismiques (naturelles ou artificielles). Elles sont ainsi limitées par le nombre, la puissance et le taux d'occurrence de ces sources. Récemment, une alternative à ces méthodes actives a été proposée, basée sur l'utilisation de champs diffus (bruit sismique ambiant ou coda de séismes). Ces méthodes passives permettent de reconstruire des fonctions de Green dans le milieu à partir de l'intercorrélation d'enregistrements entre toutes les paires d'un réseau. L'objectif de cette thèse a été d'étudier les limites de ces techniques passives, afin d'en définir les champs d'application.

Ce travail est centré sur les petites échelles de la géophysique (10 m à 1 km) et aux hautes fréquences (5 à 50 Hz), où les propriétés du champ d'onde mettent souvent en défaut l'hypothèse de champ diffus du théorème de corrélation. A partir de cas expérimentaux où la fonction de corrélation ne converge pas vers la fonction de Green, de nouvelles approches sont définies afin de retrouver malgré tout les propriétés du milieu.

L'importance de la distribution des sources de bruit pour les milieux non diffusant est mise en évidence. La caractérisation du champ d'onde est ainsi essentielle avant la mise en place de méthodes passives. Une étude comparative des méthodes d'imagerie active et passive est proposée. Elle montre que la résolution de l'imagerie passive est au moins aussi bonne que celle de l'imagerie active.

Enfin, une approche méthodologique est proposée pour l'étude du bruit ambiant, à partir d'outils issus de l'acoustique sous marine. Son application à l'étude de zones hydrothermales est discutée.

ABSTRACT

Traditional methods for ground studies are based on the recording of seismic waves from triggered natural or artificial sources. They are limited by the number, the power and the occurrence rate of those sources. Recently, an alternative method has been proposed, based on the use of diffuse wavefields (ambient seismic noise or coda waves). Those passive methods yield the Green's function in the propagation medium from the cross-correlation of recordings between every receiver pair among the array. The objective of this work is to study the advantages and limitations of passive cross-correlation techniques, in order to define their application fields.

We focused on small scale geophysics (10 m to a kilometer), and high frequencies (5 to 50 Hz), where the wavefield properties are usually incompatible with the diffuse field hypothesis of the correlation theorem. From experimental configurations in which we point out and demonstrate that the correlation function does not converge toward the Green's function, our goal is to find alternative processing methods to nevertheless measure properties of the medium.

We point out that the noise source distribution is of great importance when there is no scattering in the medium. It results that characterizing the seismic noise wavefield is mandatory prior to any cross-correlation processing. A comparative study of active and passive imaging is addressed. It shows that the resolution of passive imaging is at least as good as active imaging.

Finally, we developed a methodological approach to characterize ambient noise sources using some array processing tools taken from ocean acoustics. Application to hydrothermal areas is discussed.

Merci !

Aux membres de mon jury, et en particulier mes rapporteurs, qui ont accepté d'évaluer ce travail.

A Philippe et Michel, mes directeurs de thèse, qui ont permis l'aboutissement de cette thèse.

Leur disponibilité, leur dynamisme et leur clair-vision scientifique
ont rendu ces trois années très agréables et productives.

J'ai aussi beaucoup appris de leurs différences
de caractère et de points de vue, toujours très stimulantes.

A Gérard Herman, qui a initié une grande aventure scientifique avant de nous quitter.

A Arie Verdel et Xander Campman avec qui j'ai eu beaucoup de plaisir à travailler.

Je remercie également Shell R&D dans son ensemble, le Sultanat d'Oman
et Petroleum Development of Oman de m'avoir donné l'accès
aux données exploitées dans le second chapitre de ce manuscrit,
et de m'avoir autorisé à publier les résultats obtenus.

Aux étudiants du LGIT qui, par l'ambiance conviviale qu'ils entretiennent,
participent à la réalisation d'un travail de qualité.

Aux permanents du LGIT, non moins méritants en ce qui concerne l'ambiance,
qui savent se rendre disponibles et abordables.

A mon épouse qui a su être patiente dans les périodes de stress
inhérentes aux travaux de recherche.

A ceux qui m'ont aidé et soutenu pendant la difficile période
de la rédaction de ce manuscrit et de la préparation de la soutenance de thèse.

A ceux qui ont participé à la préparation du traditionnel pot de soutenance,
étape importante de la thèse.

Table des matières

Table des matières	7
Contexte de la thèse	11
1 Ondes, fonctions de Green et corrélations	15
1.1 Bases théoriques	16
1.1.1 Équation d'onde	16
1.1.2 Fonction de Green	16
1.1.3 Différents types d'ondes	17
1.2 La coda	17
1.2.1 Définition et propriétés	17
1.2.2 Le facteur de qualité « Coda Q »	21
1.3 Utilisation du bruit sismique	21
1.4 Corrélations et fonctions de Green	23
1.4.1 Aperçu historique	23
1.4.2 L'analogie avec le retournement temporel	25
1.4.3 Imperfections de la reconstruction	27
1.4.4 Cross-correlation of random fields : mathematical approach	28
Bibliographie chapitre 1	37
2 Convergence vers la fonction de Green	43
2.1 Description du jeu de données Qarn Alam	44
2.2 Convergence of the two-point correlation function toward the Green's function in the context of a seismic prospecting dataset	44
2.2.1 Abstract	45
2.2.2 Introduction	45
2.2.3 The seismic network	47
2.2.4 Correlation function vs source distribution	47
2.2.5 Amplitude of the crosscorrelation function	52
2.2.6 Conclusion	54

2.2.7	Acknowledgments	54
2.3	Imagerie passive d'un milieu homogène	54
2.3.1	Analyse fréquence/nombre d'onde (FK)	54
2.3.2	Application de l'analyse FK aux données Qarn Alam	57
2.4	Tomographie passive d'ondes de surface	60
2.4.1	Algorithme d'inversion	62
2.4.2	Inversion de la zone de Qarn Alam	64
2.4.3	Conclusion et perspectives sur l'imagerie passive	70
2.5	Caractérisation de la diffusion : mesure du libre parcours moyen élastique	71
	Bibliographie chapitre 2	76
3	Mesure de courbes de dispersion à partir de corrélations de bruit	81
3.1	Méthodes passives d'analyse en réseau	82
3.1.1	La méthode <i>Spatial Autocorrelation</i> (SPAC)	82
3.1.2	La méthode fréquence/nombre d'onde (FK)	82
3.2	Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique	83
3.2.1	Summary	84
3.2.2	Introduction	84
3.2.3	Ambient noise simulation and array configuration	86
3.2.4	Phase velocity dispersion curves obtained from HRFK and SPAC analysis	90
3.2.5	Phase velocity dispersion curves obtained from noise correlations slantstack (NCSS)	92
3.2.6	Results comparison	94
3.2.7	Conclusion	98
	Bibliographie chapitre 3	100
4	Utilisation de sources non synchronisées	105
4.1	Small scale geophysics inversion using surface waves extracted from noise cross-correlation	106
4.1.1	Abstract	106
4.1.2	Introduction	106
4.1.3	Retrieving a seismic section from user-defined noise cross correlation	107
4.1.4	Geophysics inversion	111
4.1.5	Conclusion	112
4.2	Étude d'un milieu latéralement inhomogène	113
4.2.1	Étude des profils linéaires	114
4.2.2	Utilisation du réseau en spirale	120
4.2.3	Conclusion sur cette étude	124
4.3	Étude de la Solfatare, Italie	124
4.3.1	Présentation de la Solfatare	124
4.3.2	Étude des données actives	126
4.4	Conclusion sur l'utilisation de sources non synchronisées	129
	Bibliographie chapitre 4	131

5	Caractérisation du bruit ambiant	133
5.1	Méthodes	134
5.1.1	Formation de voie (<i>Beamforming</i>)	134
5.1.2	Décomposition en valeurs singulières	136
5.1.3	A propos du rang de la matrice de densité spectrale croisée	137
5.1.4	Ambiguïté du modèle de vitesse	138
5.2	Étude de la vallée de Waimangu, Nouvelle-Zélande	138
5.2.1	Présentation de la vallée de Waimangu	138
5.2.2	Influence du cycle	140
5.2.3	Traitement fréquentiel	144
5.2.4	Comparaison avec la polarisation spontanée	144
5.3	Étude passive de la Solfatare, Italie	147
	Bibliographie chapitre 5	152
	Conclusion et perspectives	155
A	Cross-correlation of random fields : mathematical approach and applications	159
	Bibliographie générale	179

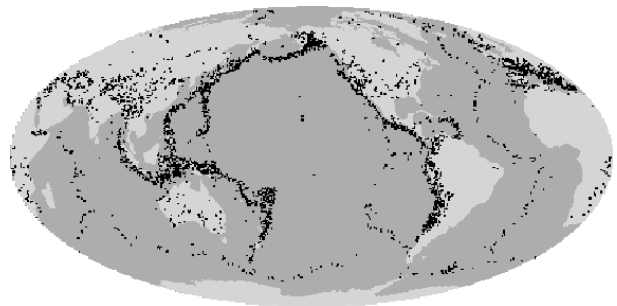
Contexte de la thèse

Les méthodes d'analyse classiques du sous sol utilisées en sismologie reposent sur l'enregistrement de séismes, ce qui entraîne deux limitations principales :

1) La répartition des séismes à la surface de globe est limitée à quelques régions géologiquement actives, principalement les limites de plaques (voir la figure 1). La résolution dans toutes les autres régions, où il n'y a pas de séisme, est alors très limitée. En particulier, lorsque les récepteurs sont situés loin des régions sismiques, les informations hautes fréquences sont perdues à cause de l'atténuation au cours de la propagation.

2) Le taux d'occurrence des séismes est faible, ce qui interdit le suivi en temps réel de l'évolution de structures actives comme les volcans ou les failles.

FIG. 1 – Distribution moyenne de la sismicité mondiale. Les séismes sont localisés principalement le long des failles actives.



A des échelles plus petites, les sources utilisées sont artificielles et dépendent de l'étendue spatiale de la zone à étudier. En voici quelques exemples, classés par ordre d'énergie de la source et donc de portée (de la dizaine de mètres à la dizaine de kilomètres) : coups de masse, canon à air comprimé, pots vibrants, explosifs, camions vibrateurs. Les fréquences mises en jeu sont alors de quelques hertzs à la centaine de hertzs. La résolution des images obtenues par la sismique active est limitée par le nombre et la puissance des sources utilisées. Il est alors difficile de faire l'image de grandes régions ou de structures profondes à cause des moyens nécessaires. De plus, les sources artificielles sont onéreuses et dures à mettre en œuvre dans les endroits d'accès difficile comme le fond des océans ou les terrains très accidentés. Le suivi temporel de structures géologiques est limité par les moyens humains et financiers, le coût de chaque acquisition étant élevé. Il est également techniquement difficile de réaliser les sources artificielles de grande énergie reproductibles dans le temps nécessaires au suivi de structures géologiques complexes comme les volcans.

L'imagerie passive est une alternative à ces méthodes actives. Elle est basée sur l'utilisation de champs diffus (bruit sismique ambiant, coda de séismes) afin de reconstruire les fonctions de Green (également appelées réponses impulsionnelles) du milieu entre différents points de mesure via l'intercorrélation des enregistrements. Les enjeux liés à cette technique sont importants puisqu'elle permet de s'affranchir de beaucoup de limitations liées à l'utilisation de sources naturelles ou artificielles. Ainsi, l'utilisation du bruit ambiant n'est pas limitée dans le temps ou dans l'espace, puisque la Terre est un environnement naturellement bruité, en tout point de sa surface et à tout moment. Une fois un réseau déployé, l'acquisition peut alors se faire sans intervention humaine, ce qui réduit considérablement les coûts de mise en œuvre. De plus, la résolution n'est plus liée au nombre de sources utilisées, mais au réseau de récepteurs. Elle est donc totalement contrôlée, même dans les régions asismiques ou les études actives sont pénalisées par l'éloignement des sources. En ce qui concerne le suivi temporel de structures, le bruit ambiant est souvent stationnaire sur de très longues périodes de temps. Les résolutions temporelles sont meilleures que celles obtenues avec des expérimentations actives, les durées de convergence étant en général plus courtes que les durées entre deux séismes ou que le temps nécessaire à la mise en place de sources artificielles répétitives.

Cette thèse a débuté alors que les méthodes de corrélation étaient en plein essor, avec un nombre croissant de publications à ce sujet. Il existe différentes approches théoriques de ces méthodes, basées sur différents jeux d'hypothèses de base. Toutes montrent que la convergence de l'intercorrélation vers la fonction de Green dépend des propriétés statistiques du champ d'onde utilisé, et que son isotropie a une importance toute particulière. L'objectif de cette thèse a été d'explorer les limites de l'imagerie passive, appliquée en particulier aux petites échelles de la géophysique (de la dizaine de mètres au kilomètre) et à la sub-surface. À ces échelles, les fréquences utilisées sont comprises entre quelques hertzs et une centaine de hertzs et les champs d'onde enregistrés sont rarement isotropes car ils sont sensibles aux hétérogénéités locales du milieu. De plus, la forte atténuation des ondes dans le sous sol terrestre pour ces fréquences implique que les sources de bruit soient locales, et donc souvent peu nombreuses, ce qui limite l'effet de moyenne sur les sources qui aide à rendre le champ isotrope aux basses fréquences. Ces caractéristiques du champ mettent en défaut les hypothèses habituellement utilisées pour démontrer le théorème de corrélation. Dans ce contexte, j'ai cherché à mieux comprendre les conséquences de ces défauts d'isotropie sur la convergence des fonctions de corrélation vers la fonction de Green. S'est alors posé la question de savoir si dans de tels cas il n'était pas possible, moyennant éventuellement des hypothèses supplémentaires, de retrouver un cas d'application des méthodes de corrélation prenant en compte les propriétés effectives du champ d'onde.

Ce manuscrit comporte 5 chapitres. Le premier constitue une introduction rappelant le contexte et les bases physiques nécessaires à la compréhension des méthodes de reconstruction de fonctions de Green par corrélation. Le deuxième chapitre est consacré à l'étude de la convergence de la fonction de corrélation vers la fonction de Green en fonction de la distribution de sources utilisée ainsi que de la diffusion dans le milieu. Afin de pouvoir contrôler l'influence de ces paramètres, tout en utilisant des données réelles, cette étude sera basée sur l'utilisation d'un jeu de données actif, mais traité comme le seraient des données passives. Dans ce chapitre seront également abordées les techniques d'imagerie sismique, à partir de jeux de données actifs ou passifs. Le chapitre 3 traitera de l'utilisation du bruit ambiant afin de mesurer des courbes de dispersion via l'utilisation de techniques de corrélation. L'utilisation de données synthétiques permettra d'évaluer la précision des

méthodes utilisées dans le cas d'un milieu complexe mais parfaitement connu. Ces données synthétiques seront également utilisées pour comparer les performances des méthodes de mesure de courbes de dispersion par corrélation à celles d'autres méthodes passives classiquement utilisées en géophysique. Les stratégies à adopter en fonction de la distribution de bruit y seront discutées. Le quatrième chapitre montrera comment traiter dans la pratique les cas où le bruit ambiant ne possède pas la propriété d'isotropie nécessaire à l'utilisation du théorème de corrélation dans son énoncé le plus simple. L'utilisation de sources non synchronisées permet alors de retrouver un cas d'application de ces méthodes. Le cas de milieux présentant des variations latérales sera aussi étudié. Enfin, le dernier chapitre sera consacré à l'étude du bruit proprement dit, et plus particulièrement à la localisation des sources de bruit, dont on aura vu qu'elle influence la convergence des fonctions de corrélation vers les fonctions de Green.

Ondes, fonctions de Green et corrélations

Sommaire

1.1 Bases théoriques	16
1.1.1 Équation d'onde	16
1.1.2 Fonction de Green	16
1.1.3 Différents types d'ondes	17
1.2 La coda	17
1.2.1 Définition et propriétés	17
1.2.2 Le facteur de qualité « Coda Q »	21
1.3 Utilisation du bruit sismique	21
1.4 Corrélations et fonctions de Green	23
1.4.1 Aperçu historique	23
1.4.2 L'analogie avec le retournement temporel	25
1.4.3 Imperfections de la reconstruction	27
1.4.4 Cross-correlation of random fields : mathematical approach	28
Bibliographie chapitre 1	37

Ce chapitre d'introduction expose les bases théoriques de la physique des ondes indispensables à la lecture de ce manuscrit. Les notions et les objets qui y seront utilisés seront également introduits. Nous aborderons ainsi les bases de la propagation des ondes dans les solides, et les différents types d'ondes élastiques seront présentés. La notion de champ diffus sera ensuite introduite, à partir de deux exemples fondamentaux : la coda sismique et le bruit sismique ambiant. Enfin, le principe des méthodes de corrélation qui font l'objet de cette thèse sera exposé à partir d'un aperçu historique et d'une analogie simple, avant d'être démontré rigoureusement à l'aide d'une théorie mathématique adaptée.

1.1 Bases théoriques

► 1.1.1 Équation d'onde

Si on considère une grandeur physique quelconque φ , scalaire ou vectorielle, liée à une excitation s , son évolution spatiale et temporelle au sein d'un milieu continu sans perte peut être régie par deux équations :

$$\text{l'équation de diffusion : } \left[\Delta - \frac{1}{D} \frac{\partial}{\partial t} \right] \varphi(\vec{r}, t) = s(\vec{r}, t)$$

$$\text{l'équation d'onde : } \left[\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \varphi(\vec{r}, t) = s(\vec{r}, t)$$

D est appelé coefficient de diffusion et c est la vitesse de l'onde représentée par φ .

Ces deux équations ont des propriétés très différentes. En particulier, l'équation d'onde est symétrique par renversement du temps, c'est-à-dire que si $\varphi(\vec{r}, t)$ est solution de l'équation, $\varphi(\vec{r}, -t)$ l'est aussi. L'équation de diffusion, elle, ne possède pas cette symétrie. Cette différence simple (mais fondamentale) a des conséquences sur la physique des processus décrits par ces équations : la propagation des ondes est réversible, tandis que la diffusion ne l'est pas. Ainsi, une onde divergente peut être rendue convergente (à l'aide d'une lentille ou d'un miroir par exemple), tandis que la diffusion d'un liquide ne peut se faire que dans le sens d'une augmentation d'entropie, et jamais dans le sens inverse. Par conséquent, les ondes peuvent être utilisées pour obtenir des informations sur leur source, et en faire une image, tandis que l'observation des phénomènes de diffusions ne permet pas de remonter à leur origine.

Dans ce manuscrit, seule l'équation d'onde sera étudiée, dans le cas particulier de la propagation d'ondes mécaniques dans les solides homogènes, isotropes et linéaires. La plupart des résultats obtenus découlent cependant directement des propriétés de l'équation d'onde, et peuvent être étendus à n'importe quel phénomène physique régi par cette équation.

► 1.1.2 Fonction de Green

Pour un terme source s fixé, il y a unicité de la solution φ de l'équation d'onde pour un jeu de conditions aux limites donné. On appelle fonction de Green (notée ici g_0 , aussi appelée réponse impulsionnelle) la solution particulière de l'équation d'onde pour une source impulsionnelle, et les conditions aux limites suivantes :

$$\left\{ \begin{array}{l} g_0(\vec{r} \rightarrow \infty) = 0 \\ g_0(t = 0) = 0 \\ \frac{\partial g_0}{\partial t}(t = 0) = 0 \end{array} \right.$$

On a alors :

$$\left[\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] g_0(\vec{r}, t) = \delta(\vec{r}) \delta(t) \quad . \quad (1.1)$$

Dans cette définition mathématique de la fonction de Green, aucune grandeur physique n'est précisée. On peut donc définir plusieurs fonctions de Green, suivant que g_0 représente un déplacement, une vitesse ou une accélération, et que la source est une impulsion en déplacement, en vitesse ou en accélération (pour ne citer que les trois grandeurs habituellement utilisées en géophysique). Ces différentes fonctions de Green étant liées entre elles par des opérations simples de dérivation, il n'est habituellement pas nécessaire de préciser les grandeurs physiques utilisées, et on parle de « la » fonction de Green.

► 1.1.3 Différents types d'ondes

Seul le cas d'ondes mécaniques se propageant dans des solides sera abordé dans ce manuscrit. La grandeur physique φ considérée dans l'équation d'onde est alors une grandeur vectorielle représentant le déplacement local de matière dû au passage de l'onde. Dans un solide isotrope infini, ce déplacement apparaît comme la somme de deux ondes, l'une polarisée longitudinalement (onde de compression ou onde P) et l'autre transversalement (onde de cisaillement ou onde S) — voir figure 1.1. Leur polarisations étant différentes, ces deux ondes de volume se propagent de façon indépendante dans le solide, l'onde P étant la plus rapide (d'où son nom, onde P primaire, par opposition à l'onde S secondaire).

Dans la pratique, les solides ne sont pas infinis, et d'autres types d'ondes peuvent apparaître à ses frontières. On parle alors d'ondes de surface, leur amplitude décroissant exponentiellement dans la direction perpendiculaire à la surface. Il en existe deux types :

- **les ondes de Rayleigh** sont polarisées dans le plan perpendiculaire à la surface incluant la direction de propagation (plan sagittal) et possèdent donc une composante longitudinale et une composante transversale.
- **les ondes de Love** n'existent que si le milieu n'est pas homogène mais constitué (au moins localement) de plusieurs couches. Elles sont polarisées transversalement dans le plan parallèle à la surface.

Ces ondes de surface résultent d'interactions entre les ondes P et S. Elles sont généralement dispersives, c'est-à-dire que leur vitesse dépend de la fréquence. Leur énergie décroît moins rapidement avec la distance que celle des ondes de volume, car elles ont une expansion géométrique cylindrique et non sphérique. Elles dominent donc généralement les enregistrements sismiques dès lors que le récepteur est situé à plus de quelques longueurs d'onde de la source.

1.2 La coda

► 1.2.1 Définition et propriétés

Les différentes ondes présentées dans le paragraphe précédent ont des vitesses de propagation différentes et peuvent être identifiées sur les enregistrements sismiques, comme le montre la figure 1.2. Sur cette figure, le signal perdure plusieurs minutes après l'arrivée des ondes identifiées, alors que la source sismique ne dure que quelques secondes. Ce temps correspond à plus de

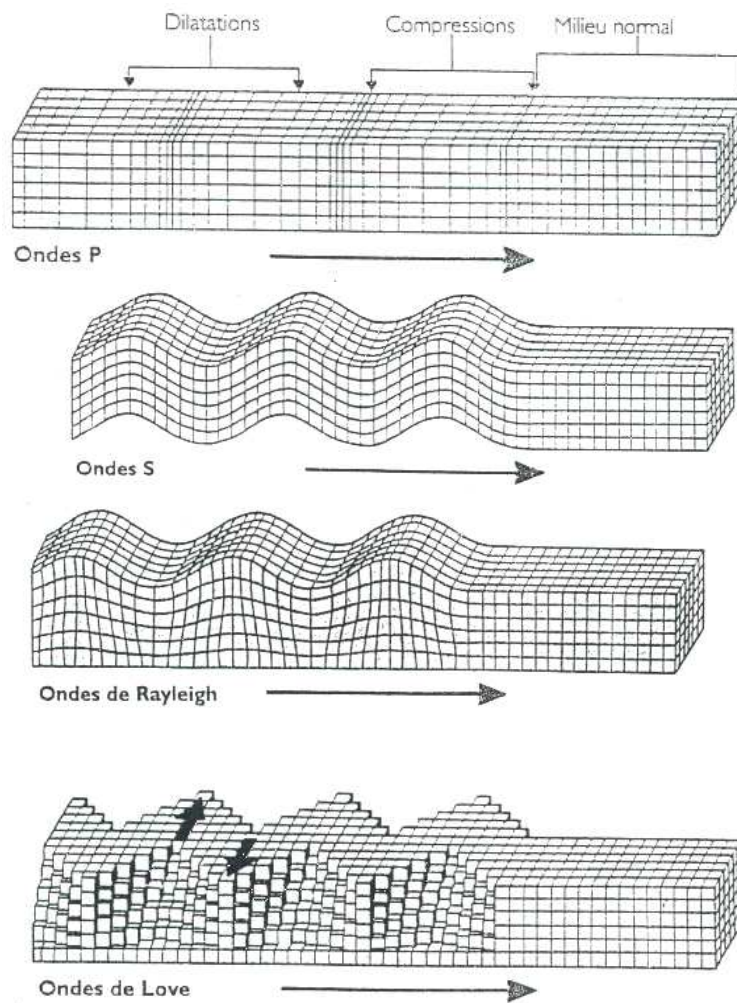


FIG. 1.1 – Les différents types d'ondes se propageant dans les solides, ainsi que les déplacements associés. Les ondes P et S sont des ondes de volume tandis que les ondes de Rayleigh et de Love sont des ondes de surface. Les flèches horizontales indiquent la direction de propagation de l'onde.

dix fois le temps de trajet des ondes directes entre la source et le sismomètre. Cette partie du signal est appelée *coda*, par analogie avec le terme désignant la fin d'un morceau de musique.

La coda a pour origine la diffusion des ondes dans le sous-sol [Aki, 1969, Aki and Chouet, 1975, Margerin et al., 1999]. Dans un milieu diffusant, seule une fraction des ondes suit un trajet balistique (ou direct), le reste interagissant avec les nombreux diffuseurs présents sur le trajet de l'onde. Les ondes ainsi diffusées parcourent plus de distance que les ondes directes avant d'arriver au récepteur et arrivent donc plus tard, ce qui explique la durée temporelle de la coda.

Dans la coda, la longueur du trajet parcouru par les ondes est liée au temps d'enregistrement. Ainsi, si les ondes ont une vitesse c , les ondes constituant la coda au temps t ont parcouru une distance ct . Les durées de coda enregistrées sont telles que les distances parcourues sont bien supérieures aux distances sources-récepteurs. On peut alors considérer que les ondes ont « sondé » un volume du milieu de propagation lié à la distance qu'elles y ont parcouru, et qu'elles contiennent ainsi une information statistique sur ce milieu. Il y a ainsi une équivalence temps d'observation/distance parcourue dans la coda : étudier la coda autour d'un temps t permet d'étudier les propriétés statistiques du milieu de propagation à une échelle qui croît avec t (et qui dépend du caractère diffusant de ce milieu).

La coda semble être un bruit sismique (et a longtemps été considéré comme tel) car sa forme d'onde est aléatoire. En effet les diffusions successives déforment les ondes et la forme du signal enregistré ne peut être prédite. La coda reste cependant parfaitement déterministe et deux enregistrements successifs réalisés au même endroit, avec la même source et le même récepteur, donneront strictement le même signal (au bruit près). Elle a un caractère chaotique dans le sens où lorsqu'un seul des diffuseurs est modifié, le signal enregistré est fortement perturbé, et ce d'autant plus que l'on est tard dans l'enregistrement [Tourin et al., 2001]. La coda est ainsi très sensible aux variations du milieu car les ondes multidiffusées effectuent un trajet dans la zone modifiée d'autant plus long que l'on est tard dans la coda. Cette propriété est exploitée en acoustique dans les techniques de DAWS (*Diffusing Acoustic Wave Spectroscopy*, [Cowan et al., 2000]) où elle permet de mesurer des infimes perturbations d'un milieu (comme par exemple le déplacement d'un seul diffuseur), ce qui ne serait pas possible avec les ondes directes. En sismologie, des méthodes similaires ont été développées, appelées méthodes des doublets [Poupinet et al., 1984, Roberts et al., 1992], ou *coda wave interferometry* [Snieder, 2002, 2006].

Les ondes de la coda possèdent des propriétés statistiques particulières : au cours de chacune des diffusions, il y a conversion de l'énergie entre les différents modes de propagation des ondes. Ainsi, au bout d'un grand nombre de diffusions, l'énergie est équirépartie entre tous les modes de propagation des ondes [Campillo et al., 1999, Margerin et al., 2000, Shapiro et al., 2000, Hennino et al., 2001]. De même, la direction des ondes change de façon aléatoire à chaque diffusion et le champ enregistré au niveau du capteur est isotrope au bout d'un temps assez long (qui dépend de la vitesse des ondes et de la diffusion dans le milieu). Ces deux régimes d'équipartition en énergie et en direction ne sont atteints qu'asymptotiquement, avec des temps caractéristiques différents [Paul et al., 2005]. Une propriété remarquable de la coda est que son intensité suit une équation de diffusion, et non une équation d'onde [Margerin et al., 1999].

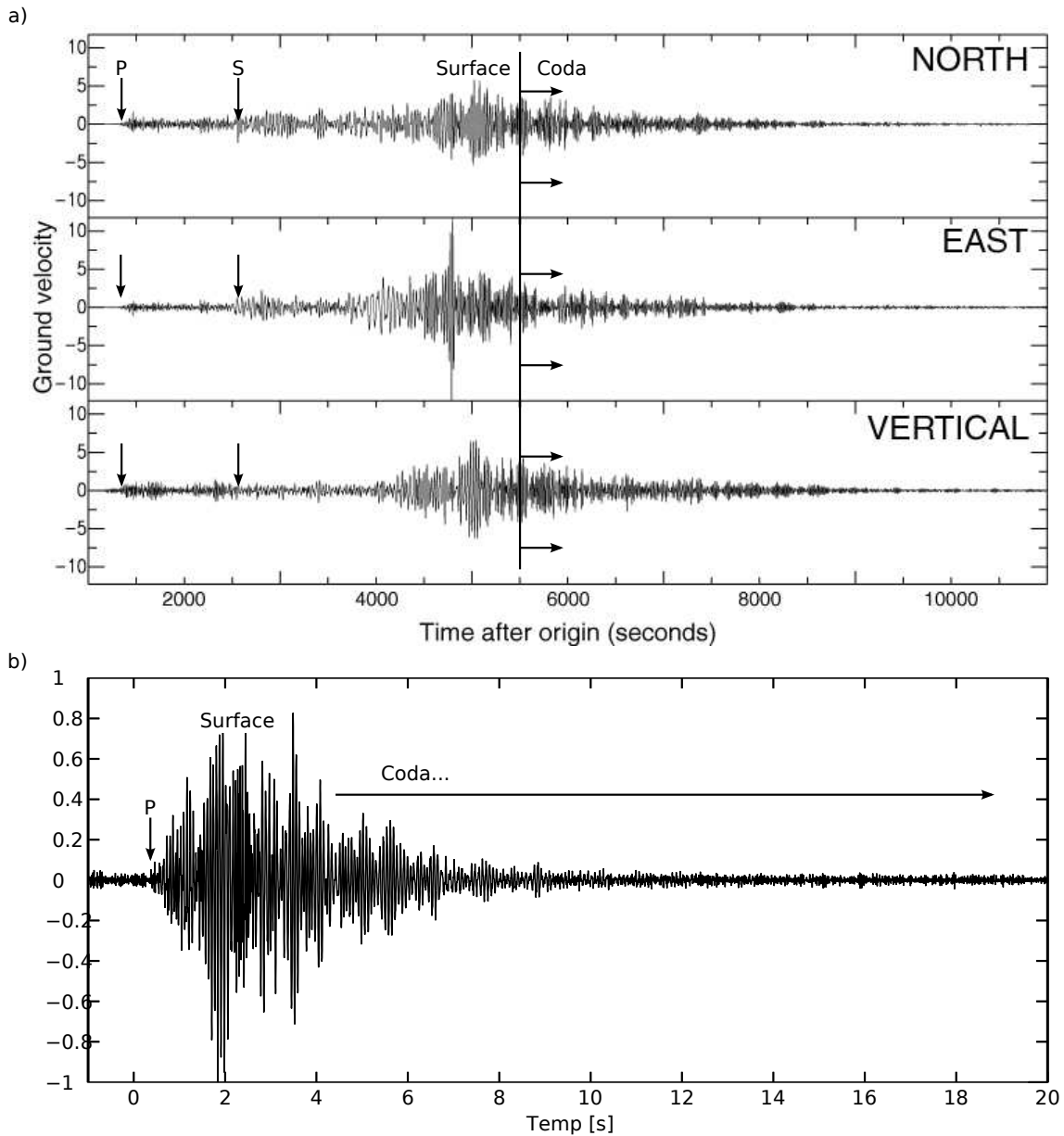


FIG. 1.2 – Exemples de sismogrammes : **a.** Enregistrement du séisme de Sumatra du 26/12/04 par une station sismique située en Arizona. On peut identifier les ondes P et S ainsi que les ondes de surface. La partie du signal qui vient ensuite est la coda. Elle dure plusieurs minutes. **b.** Sismogramme (composante verticale) enregistré lors d’une campagne d’acquisition sur le Puy des Goules (Auvergne). La source est un *sweep* engendré par un camion vibreur et est située à 745 m du géophone. Le temps de propagation des ondes de surface est d’environ 2 secondes tandis que la coda dure au moins 8 secondes. La distance parcourue par les ondes arrivant à 10 secondes est 5 fois supérieure au trajet source–récepteur.

► 1.2.2 Le facteur de qualité « Coda Q »

La coda issue de séismes, qui représente une grande partie des enregistrements, et souvent la majorité de l'énergie reçue, a longtemps été négligée dans l'analyse des signaux sismiques. Son apparence chaotique lui a valu d'être considéré comme un bruit ne portant aucune information. C'est en 1975 que Aki et Chouet se sont intéressés à ces ondes et ont montré qu'elles étaient porteuses d'information [Aki and Chouet, 1975]. Ces pionniers ont montré que la décroissance temporelle de l'énergie de la coda est une caractéristique régionale et qu'elle est indépendante des propriétés de la source (la magnitude du séisme, sa localisation, son mécanisme...). Ils montrent alors que l'enveloppe E_ω de la coda à une pulsation donnée ω suit une loi exponentielle qui peut s'écrire :

$$E_\omega(t) \propto \frac{1}{t^\gamma} \exp\left(-\frac{\omega t}{Q_c(\omega)}\right) . \quad (1.2)$$

Le facteur $1/t^\gamma$ est un facteur géométrique, γ étant compris entre 1 et 2, et Q_c est appelé *facteur de qualité de la coda*. C'est un nombre sans dimension qui dépend de la fréquence et de la région. Deux explications possibles ont alors été mises en avant [Aki, 1969, Aki and Chouet, 1975] : la coda peut être due à la diffusion simple, et la décroissance de son enveloppe liée à l'absorption dans le milieu ; elle peut aussi résulter de la diffusion multiple, et son évolution temporelle est alors entièrement expliquée par les propriétés de diffusion du milieu.

Ce n'est que 20 ans plus tard que des modèles plus complexes ont été proposés, associés à des simulations numériques, afin de décrire plus précisément la coda [Ryzhik et al., 1996, Margerin et al., 1999]. Il est désormais acquis que la coda a pour origine la diffusion multiple, et que l'évolution temporelle de son énergie est correctement décrite par l'équation du transfert radiatif [Ryzhik et al., 1996, Margerin et al., 1999].

1.3 Utilisation du bruit sismique

Comme nous l'avons vu, les études sismologiques se sont d'abord basées sur l'identification et l'analyse des ondes directes, avant de s'intéresser aux ondes diffuses et à la coda. La durée des enregistrements nécessaires à ces études a alors augmenté, ce qui a été rendu possible par l'évolution des moyens techniques d'acquisition, de traitement et de stockage des données. Les difficultés à établir des catalogues d'événements exhaustifs, ainsi qu'à définir une durée d'enregistrement utile pour la coda (celle-ci dépendant de la diffusion dans le milieu et donc de la région) ont fait évoluer les systèmes d'acquisition du mode déclenché au mode continu : on ne cherche plus de nos jours à limiter les enregistrements à de petites fenêtres temporelles incluant des événements sismiques, mais plutôt à avoir des stations d'acquisition opérationnelles en continu afin de pouvoir rechercher ces événements ultérieurement. D'immenses quantités d'enregistrements continus sont ainsi disponibles à travers des bases de données mondiales ou régionales. Ces bases de données sont revisitées régulièrement à la lumière des nouvelles techniques de détection d'événements afin de produire des catalogues toujours plus complets.

Si on extrait de ces bases de données 1 jour de signal pour une station quelconque, la probabilité pour qu'il y ait un événement sismique dans cet intervalle de temps est faible. La figure 1.3 montre

que même alors les ondes directes ou les signaux de coda, ne représentent qu'une infime partie de l'enregistrement.

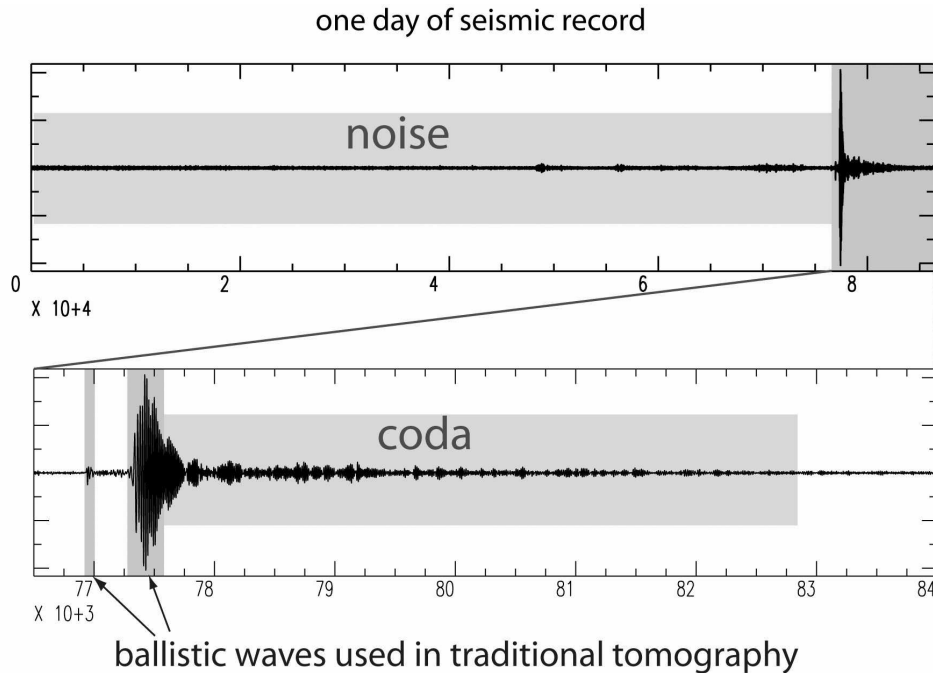


FIG. 1.3 – Un jour d'enregistrement continu incluant un séisme de grande magnitude. Environ 95 % de l'enregistrement est du bruit sismique, 4 % de la coda et les ondes directes utilisées pour l'imagerie sismique traditionnelle ne représentent qu'un pourcent du signal.

Il a alors été proposé [Shapiro and Campillo, 2004] d'utiliser le bruit sismique ambiant en le considérant comme un champ diffus. Cette hypothèse repose à la fois sur le fait que la croûte terrestre est diffusante, mais aussi sur le caractère aléatoire des sources de bruit, aussi bien en ce qui concerne leur répartition que de leur mécanisme. C'est bien sûr une hypothèse forte, dont la validité doit être vérifiée avant toute étude.

L'origine du bruit sismique dépend de la fréquence considérée. On définit en général 3 bandes de fréquences reliées à 3 phénomènes différents à l'origine bruit sismique :

- en dessous de 1 Hz, le bruit provient essentiellement des océans (couplage vagues/croûte terrestre) et des phénomènes météorologiques de grande ampleur comme les cyclones. Le bruit ambiant très basses fréquences est appelé *hum*. Il permet entre autre d'observer les modes propres de la Terre.
- autour de 1 Hz, il est dû à des phénomènes météorologiques locaux, comme le vent par exemple.
- au dessus de 1 Hz, le bruit est généralement lié à l'activité humaine : trafic routier, usines, stations de pompage...

Ces phénomènes sont d'autant plus locaux que la fréquence augmente, à cause de la forte atténuation des hautes fréquences dans le sous-sol terrestre. Ainsi, le bruit généré à très basses fréquences, avec par exemples les pics microcosmiques à 7 et 14 secondes de période, peut être enregistré en tout point du globe, même s'il n'est généré que dans les océans profonds [Stehly et al., 2006]. Inversement, le bruit au delà de 100 Hz n'est présent qu'à proximité des villes ou des zones d'activité

humaine. L'hypothèse d'isotropie du bruit ambiant n'a alors pas la même portée lorsqu'elle est invoquée à hautes ou basses fréquences. Le bruit sismique est principalement constitué d'ondes de surface, les sources étant localisées à la surface terrestre.

Les avantages à utiliser le bruit sismique ambiant sont multiples, et dépendent des échelles spatiales considérées. À l'échelle régionale, les sources utilisées pour les méthodes actives sont les séismes. Leur occurrence est limitée en temps et en espace. Ainsi, il est difficile d'étudier les régions asismiques à cause de l'éloignement des sources, et même dans les régions sismiquement actives (zones de faille, bordures de plaques continentales...) le temps d'attente avant l'occurrence d'une source est *a priori* inconnu. Ces caractéristiques rendent les études sismologiques difficiles et coûteuses, puisque les réseaux doivent parfois rester déployés (et opérationnels) sur de longues périodes. L'utilisation du bruit permet de s'affranchir de ces limitations puisqu'à ces échelles seul le bruit basse fréquence est utile, et celui-ci est présent partout sur terre. De plus la durée d'acquisition nécessaire à l'imagerie passive peut être connue à l'avance par des analyses de convergence. Un deuxième intérêt des méthodes passives est que la résolution est gouvernée par le réseau utilisé et non par la distribution de sources comme dans le cas des méthodes actives. Shapiro et al. [2005] ont ainsi pu obtenir une image tomographique de la Californie à partir d'un mois d'enregistrements continus, avec une résolution meilleure que celles obtenues par des méthodes actives utilisant plusieurs décennies d'enregistrements sismiques.

Mes travaux de thèse m'ont amené à utiliser des données autres que du bruit ambiant à proprement parler. Cependant, tous les résultats présentés dans ce manuscrit ont été obtenus grâce à des méthodes identiques à celles utilisées pour le traitement du bruit sismique. Ces travaux sont effectués aux petites échelles de la géophysique (de la dizaine de mètres au kilomètre), et donc à hautes fréquences (principalement entre 5 et 50 Hz), là où le bruit ambiant vérifie le moins l'hypothèse de champ diffus isotrope. La motivation première de ces travaux est ainsi de déterminer ce qui peut être fait dans les cas où certaines hypothèses des théories relatives à l'utilisation du bruit ne sont pas vérifiées.

1.4 Corrélations et fonctions de Green

L'intercorrélation d'enregistrements d'un champ diffus entre deux points converge vers la fonction de Green complète du milieu entre ces deux points, incluant toutes les réflexions, diffractions et modes de propagation [Weaver, 2005].

Pour démontrer ce résultat, et pour préciser sous quelles hypothèses il est valide, de nombreuses études expérimentales, numériques et théoriques ont été menées. Je propose ici d'en faire un rappel historique et d'en exposer les principaux résultats.

► 1.4.1 Aperçu historique

C'est dans le domaine de l'héliosismologie que la corrélation de bruit ambiant a été utilisée en premier : des enregistrements du mouvement de la surface du soleil ont été corrélés pour mesurer les vitesses de propagation à la surface [Duvall et al., 1993, Gilles et al., 1997]. L'idée de *daylight imaging* a été proposée dans le domaine de la prospection sismique par Claerbout [1968], qui

compare le bruit sismique à la lumière du jour : pourquoi avoir besoin d'une source d'ondes artificielle quand le milieu en contient naturellement ? Plus récemment, un article pionnier de Weaver and Lobkis [2001] montrait comment le mouvement brownien enregistré par deux transducteurs ultrasonores collés à la surface d'un bloc d'aluminium pouvait permettre de retrouver la fonction de Green complète entre ces deux points. Dans cet article, les auteurs interprètent théoriquement la reconstruction de la fonction de Green en invoquant l'équipartition des modes acoustiques excités dans le bloc d'aluminium. Ce résultat a été généralisé au cas où le caractère aléatoire du champ d'onde n'est pas issu de la répartition des sources, mais résulte de la diffusion multiple dans un milieu hétérogène [Lobkis and Weaver, 2001]. L'utilisation d'une représentation spectrale [Lobkis and Weaver, 2001], du théorème de fluctuation-dissipation [Kubo, 1966, Weaver and Lobkis, 2001, 2003, van Tiggelen, 2003, Godin, 2007], ou d'un théorème de représentation [par exemple Wapenaar, 2004] constituent des approches théoriques très rigoureuses qui permettent d'interpréter les résultats expérimentaux.

La reconstruction des fonctions de Green n'est pas une fin en soi : elles représentent ce qu'on aurait pu enregistrer à l'aide d'une expérience active, mais ne donnent pas directement d'information sur le milieu. L'étape de reconstruction des fonctions de Green n'est donc qu'une étape préliminaire à un traitement habituellement utilisé avec des données actives, et qui dépend de l'objet de l'étude. Une des applications classiques, quelque soit l'échelle et le type d'ondes considérés, est l'imagerie. Lorsque l'on parle d'« imagerie passive », il faut donc comprendre qu'on réalise deux étapes indépendantes : l'obtention des fonctions de Green entre les éléments d'un réseau, puis l'application d'un algorithme d'inversion permettant de retrouver l'image du milieu. Cette deuxième étape est identique à celle d'une « imagerie active », et peut en inclure tous les raffinements. Des études expérimentales ont démontré la faisabilité de l'imagerie passive en 1) acoustique [Lobkis and Weaver, 2001, Weaver and Lobkis, 2001, Larose et al., 2004], 2) sismologie, où Campillo and Paul [2003] ont reconstruit la fonction de Green entre deux stations sismiques à partir d'enregistrements de séismes, et 3) en océanographie en eau peu profonde où les fronts d'ondes directes et réfléchies ont été retrouvés à partir de l'intercorrélation de bruit ambiant [Roux and Kuperman, 2004, Sabra et al., 2005b]. Wapenaar [2004] a montré que, dans le cas de récepteurs situés à la surface libre d'un solide, les fonctions de Green causale et acausale étaient bien reconstruites par corrélation. Les cas de milieux non réciproques (en présence d'un flux par exemple) ou inélastiques ont aussi été étudiés [Wapenaar, 2006, Godin, 2007].

Derode et al. [2003a,b] proposent d'interpréter la reconstruction de la fonction de Green via une analogie avec le retournement temporel (voir le paragraphe 1.4.2). Ils montrent que la corrélation d'ondes diffuses peut être utilisée pour faire de l'imagerie passive en acoustique. La convergence de la corrélation de bruit vers la fonction de Green dans un milieu infini peut aussi être interprétée en utilisant le théorème de la phase stationnaire [Snieder, 2004, Roux et al., 2005b].

En sismologie, Aki [1957] avait proposé il y a longtemps d'utiliser le bruit sismique pour retrouver les propriétés de dispersion des ondes de surface dans le milieu de propagation, dans des travaux qui ont servi de base aux méthodes SPAC (voir la paragraphe 3.1.1) encore utilisées actuellement. Shapiro and Campillo [2004] ont reconstruit la partie de la fonction de Green relative aux ondes de surface en corrélant du bruit sismique entre des stations séparées de distances allant de la centaine à des milliers de kilomètres. Ils ont alors mesuré les courbes de dispersion associées à ces ondes pour des périodes de 5 à 150 secondes. Cette méthode a abouti à la première

application de l'imagerie sismique passive en Californie [Shapiro et al., 2005, Sabra et al., 2005a], avec une résolution spatiale bien meilleure que celle des techniques active habituelles basées sur les séismes. Larose et al. [2005] ont aussi utilisé la corrélation de bruit sur de petites distances sur la Lune, montrant ainsi que l'imagerie passive est assez robuste pour pouvoir être appliquée dans des environnements difficiles ou la qualité des signaux enregistrés est mauvaise.

En ce qui concerne le problème des ondes élastiques, [Sánchez-Sesma et al., 2006, 2008] ont montré que la convergence de la corrélation de bruit vers la fonction de Green reposait sur la condition d'équipartition des différentes composantes du champ élastique. En d'autres termes, l'émergence de la fonction de Green n'est possible qu'après un processus de moyenne dû à la répartition aléatoires des sources de bruit (ou des diffuseurs) lorsque l'on considère des enregistrements longs ou de la diffusion [Campillo, 2006, Larose et al., 2006].

L'objectif des techniques de corrélation est ainsi de produire des données à partir de champs diffus. En effet, elles ne permettent pas directement de mesurer une courbe de dispersion, ou de faire l'image d'un milieu. Quand on parle d'imagerie passive par corrélation, les corrélations ne sont qu'une première étape permettant la reconstruction des fonctions de Green entre les capteurs. Une technique d'imagerie « classique » est ensuite appliquée à ces fonctions de Green pour retrouver la structure du sous sol. Cette deuxième étape peut être réalisée en utilisant n'importe laquelle des techniques d'imagerie « actives ».

Deux « familles » de champs diffus sont en général considérées dans les différentes études théoriques :

- on peut considérer un milieu homogène et des sources réparties et excitées de façon aléatoire. L'équation de propagation pour chacune des sources prises séparément est simple et le champ d'onde est composé uniquement d'ondes directes, sa complexité venant de la multitude de sources.
- on peut aussi obtenir un champ diffus à partir d'une unique source, placée dans un milieu diffusant (c'est le cas de la coda). La complexité vient alors de l'équation de propagation, la contribution de chaque diffuseur étant à prendre en compte. On utilise dans ce cas des représentations statistiques pour caractériser le champ d'onde, telle que l'équation du transfert radiatif [Ryzhik et al., 1996, Margerin et al., 1999].

La difficulté dans la représentation théorique des corrélations de bruit ambiant est que le champ d'onde mélange des ondes directes et des ondes diffusées. Deux approches statistiques doivent alors être menées en parallèle. Les travaux de Colin de Verdière [2008], exposés dans le paragraphe 1.4.4, permettent de s'affranchir de cette difficulté en gardant une équation de propagation très générale.

► 1.4.2 L'analogie avec le retournement temporel

Comme nous venons de le voir, de nombreuses démonstrations du principe de reconstruction de la fonction de Green ont été proposées. Elles se basent sur des hypothèses différentes (milieu fini ou infini, homogène ou hétérogène, absorbant ou non...) et sont plus ou moins faciles à lire. Avant d'exposer dans le paragraphe 1.4.4 une démonstration rigoureuse et la plus générale possible, j'ai voulu présenter ici une approche à mon sens plus facile à appréhender : celle du retournement temporel, développée par Derode et al. [2003a,b].

Si on considère une source située en un point S émettant un pulse spatio-temporel, les signaux enregistrés en deux points A et B de l'espace sont les réponses impulsionnelles $h_{SA}(t)$ et $h_{SB}(t)$. La corrélation (notée \times) entre ces deux signaux s'écrit

$$h_{SA} \times h_{SB}(t) = \int_{-\infty}^{\infty} h_{SA}(t + \tau) h_{SB}(\tau) d\tau \quad . \quad (1.3)$$

La corrélation peut s'écrire en fonction de la convolution (notée \otimes) comme

$$h_{SA}(t) \times h_{SB}(t) = h_{SA}(-t) \otimes h_{SB}(t) \quad . \quad (1.4)$$

Si le milieu est réciproque (dans la pratique, si le milieu est au repos), $h_{SA}(-t) = h_{AS}(-t)$ et l'équation précédente peut aussi s'écrire

$$h_{SA}(t) \times h_{SB}(t) = h_{AS}(-t) \otimes h_{SB}(t) \quad . \quad (1.5)$$

La forme obtenue est celle qu'on enregistrerait dans l'expérience fictive suivante : une impulsion est émise en A , enregistrée en S où le signal est retourné temporellement et réémis pour être finalement enregistré en B (cf figure 1.4).

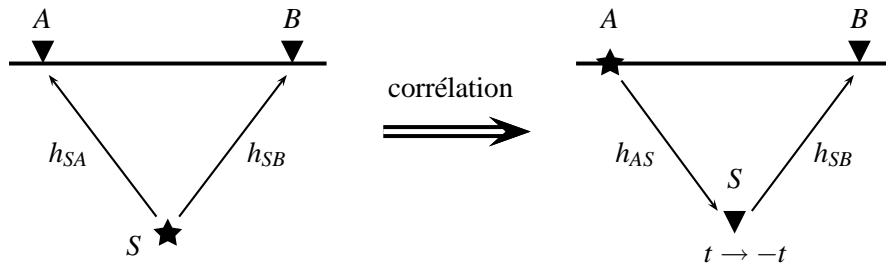


FIG. 1.4 – Corréler les signaux reçus en A et B revient à réaliser l'expérience fictive qui consiste à émettre un pulse en A , l'enregistrer en S , le retourner temporellement et le réémettre pour l'enregistrer en B .

Cette analogie permet d'appliquer les résultats du retournement temporel aux corrélations. Dans une expérience de retournement temporel, il faut enregistrer le champ en différents points entourant parfaitement la source (le miroir à retournement temporel). La refocalisation est alors possible en réémettant les signaux enregistrés après les avoir retournés dans le temps. On obtient ainsi une cavité de retournement temporel sans perte d'énergie. C'est alors la somme de toutes les ondes réémises en chaque point du miroir qui va permettre de recréer le front d'onde convergeant vers la source (voir Fink [1992], Cassereau and Fink [1992], Fink et al. [2000]). Après avoir focalisé au point source, l'onde continue de se propager en un front divergeant.

Si on applique ce principe à la corrélation, on obtient le résultat suivant : on peut retrouver la fonction de Green h_{AB} en faisant la somme des corrélations entre les signaux reçus en A et en B pour différentes sources S entourant parfaitement les points A et B . Le résultat obtenu est alors la somme des fonctions de Green causale et acausale (qui correspondent aux ondes divergente et convergente de l'expérience de retournement temporel) :

$$\frac{d}{dt} \sum_{S_i} h_{S_i A} \times h_{S_i B}(t) = h_{AB}(-t) - h_{AB}(t) \quad . \quad (1.6)$$

► 1.4.3 Imperfections de la reconstruction

Dans ce paragraphe, nous allons considérer le cas de la corrélation d'enregistrements de bruit sismique entre deux points A et B. Les remarques faites ici s'appliquent aux cas autres que le bruit, mais considérer ce cas particulier permet de fixer les idées. On prendra de plus comme convention que les temps positifs de la corrélation correspondent à la propagation de A vers B, et les temps négatifs à la propagation de B vers A.

Le qualité de la reconstruction de la fonction de Green par les corrélations dépend des propriétés du bruit utilisé. Ainsi, le bruit sismique étant généralement généré et enregistré à la surface, et les ondes de surface étant les plus énergétiques dans les enregistrements, c'est principalement la partie de la fonction de Green relative à ces ondes qui est reconstruite. Bien que des ondes de volume aient déjà été observées dans les corrélations [Roux et al., 2005a], elles n'ont jamais été utilisées à cause de la difficulté à les exploiter (en particulier à cause de leur faible rapport signal sur bruit).

La corrélation étant une opération linéaire, le spectre de la fonction de Green reconstruite ne peut inclure des fréquences qui ne sont pas présentes dans les enregistrements initiaux. Ce n'est donc qu'une version filtrée de la fonction de Green qui est reconstruite, la bande de fréquence à considérer étant celle des enregistrements (elle dépend donc à la fois du champ d'onde et du matériel d'acquisition).

Les fonctions de Green causale et acausale sont reconstruites par corrélation. Dans un milieu réciproque, ces deux parties sont identiques, et la dérivée de la corrélation devrait donc être antisymétrique (voir plus loin la relation (1.20)). Dans la pratique, cette propriété de symétrie des corrélations est rarement respectée, en particulier en ce qui concerne leur amplitude. Cette asymétrie peut avoir plusieurs origines :

- si le bruit n'est pas parfaitement isotrope, mais possède une direction de propagation privilégiée, par exemple de A vers B, alors la corrélation aura une amplitude plus importante dans les temps positifs.
- les sources de bruit n'étant pas toutes de même nature, on rencontre parfois des cas où les sources de bruit situées de part et d'autre des récepteurs n'ont pas le même spectre. Dans la corrélation, le spectre des temps positifs ne sera alors pas le même que celui des temps négatifs, ce qui peut introduire une asymétrie de l'amplitude mais également de la phase.

Dans la pratique, la fonction de corrélation présente des écarts à la fonction de Green théorique, en particulier au niveau de son amplitude et de son spectre. Ces erreurs sont liées aux propriétés du champ d'onde, qui n'est jamais parfaitement isotrope, équipartitionné et à spectre infini, comme on le suppose dans les démonstrations théoriques. Lorsque l'on veut utiliser les techniques d'imagerie passive basées sur la corrélation, il faut donc avant tout étudier les propriétés du champ d'onde enregistré afin de pouvoir prévoir et estimer les défauts dans la convergence vers la fonction de Green.

Les asymétries de la fonction de corrélation sont souvent utilisées comme indication sur la qualité de la reconstruction de la fonction de Green par corrélation. Ainsi, lorsqu'on souhaite réaliser une tomographie passive par exemple, il est nécessaire de calculer de nombreuses fonctions de corrélation, qu'il faut ensuite trier afin de minimiser les erreurs de mesure des temps de parcours. La symétrie des corrélations est alors un critère simple permettant d'éliminer les fonctions de corréla-

tion n'ayant *a priori* pas convergé vers la fonction de Green (voir par exemple [Shapiro et al., 2005, Bensen et al., 2007]).

Une application des méthodes de corrélation est de mesurer les propriétés du bruit sismique (et non du milieu) grâce à la connaissance des écarts entre les corrélations et la fonction de Green réelle. Cela ne nécessite pas forcément de connaître cette dernière, certaines de ses propriétés étant connues, en particulier les propriétés de symétrie que l'on vient d'évoquer. C'est ce qu'ont fait Stehly et al. [2006] en interprétant les défauts de symétrie en amplitude afin de retrouver les directions des flux d'énergie du bruit sismique dans les basses fréquences, et en en déduisant son origine géographique. Les défauts de symétrie en phase de la fonction de Green reconstruite peuvent aussi être analysés et attribués à des causes physiques ou non (anisotropie du bruit, défauts de l'instrumentation, et en particulier des horloges), comme l'ont montré Stehly et al. [2007].

► 1.4.4 Cross-correlation of random fields: mathematical approach

Ce paragraphe présente un extrait de l'article de revue intitulé « Cross-correlation of random fields: mathematical approach and applications » [Gouédard, Stehly, Brenguier, Campillo, Colin de Verdière, Larose, Margerin, Roux, Sánchez-Sesma, Shapiro, and Weaver, 2008] publié au cours de ma thèse dans une édition du journal *Geophysical Prospecting* dont le sujet est *What can EP learn from seismology and vice-versa*. L'article complet est présenté en annexe. Je présente ici une démonstration mathématique rigoureuse du résultat exposé précédemment sur la corrélation de bruit, dans laquelle aucune supposition n'est faite à propos du milieu de propagation. En effet l'équation de propagation utilisée ici a une forme très générale et inclut les milieux atténuants. Le cas d'un bruit blanc est d'abord abordé, puis celui plus général d'un champ d'onde quelconque vérifiant l'équipartition aux frontières du milieu (celles-ci pouvant être repoussées à l'infini).

Avant de présenter la démonstration complète, je vais en résumer ici les étapes clé. Le lecteur désireux de suivre plus en détail le déroulement de la démonstration pourra se référer à la retranscription de l'article décrivant tout le formalisme mathématique. Les notations et les numérotations des équations seront celles de la démonstration complète.

Dans le cas d'un bruit blanc, on commence par définir le noyau intégral d'un opérateur P , noté $\llbracket P \rrbracket$ (équation 1.8), qui est une extensions de la notion de matrice d'un opérateur aux cas des espaces vectoriels de base continue. On montre alors que la fonction de Green G_a peut être définie comme le noyau intégral d'un opérateur particulier (équation 1.16). L'expression de la dérivée de la fonction de corrélation est ensuite exprimée en fonction du même noyau intégral, en profitant de la propriété particulière du bruit blanc. L'équation finale est la suivante :

$$\frac{d}{d\tau} C(\tau, \vec{r}_A, \vec{r}_B) = \frac{-\sigma^2}{4a} (G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)) \quad . \quad (1.20)$$

L'hypothèse du bruit blanc est très forte, et peu adaptée aux cas pratiques. On peut s'en affranchir dans le cas d'un milieu diffusant, en la remplaçant par une hypothèse moins forte : celle de l'équipartition des modes de propagation aux frontières du milieu seulement. Cette démonstration est réalisée pour un espace de dimension quelconque d . La démonstration passe par la définition

du projecteur spectral (équation 1.22), qui est l'extension de la décomposition modale dans le cas de modes continus. On montre alors que les fonctions de corrélation et de Green peuvent toutes deux s'écrire en fonction du projecteur spectral. Le résultat final est présenté dans le domaine de Fourier :

$$\hat{C}(\omega, \vec{r}_1, \vec{r}_2) = - \frac{2^{d+1} \pi^{d-1}}{\sigma_{d-1}} \frac{c^d}{\omega^{d-2}} \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) \quad , \quad (1.36)$$

où σ_{d-1} est un facteur de normalisation qui dépend de la dimension de l'espace, \Im représente la partie imaginaire, et c la vitesse des ondes dans le milieu. Dans le cas d'un espace à 3 dimensions, un facteur $1/\omega$ apparaît et on retrouve bien le résultat de l'équation (1.20). L'hypothèse de l'équipartition des modes aux frontières du milieu est ici indispensable pour pouvoir écrire l'équation (1.27) dans laquelle le poids relatif de chaque mode est inclus.

Voici la démonstration complète telle qu'elle a été publiée dans [Gouédard et al., 2008].

The case of homogeneously distributed white noise sources

The scope of this section is to summarize the different theoretical approaches using mathematical tools that allow a global view of the correlation problem in any propagation medium. We will see that cross-correlation of noise recorded at two distant stations A and B yields the Green's function, assuming that the wavefield is a white noise distributed everywhere in the medium, with no assumption about the medium.

We consider any medium X , that does not need to be homogeneous, where the wave propagation equation is controlled by a damped equation that can be written as:

$$\frac{\partial^2 u}{\partial t^2} + 2a \frac{\partial u}{\partial t} - Lu = f \quad (1.7)$$

Here $a > 0$ is a constant that corresponds to the attenuation of the medium, $f(t, \vec{r})$ is the source field (*i.e.*, the noise field in our case) and $u(t, \vec{r})$ denotes the displacement field. If $L = c^2(\vec{r}) \Delta$, we recognize the usual wave equation. In a more general calculation, L can be any negative self-adjoint elliptic differential operator. In more physical terms, L is an operator which preserves energy.

First of all we will introduce a definition of the Green's function in the frequency domain using the *integral kernel* of the operator L , and show that this definition is equivalent to the usual one. Then, by expressing the displacement field using the Green's function, we will calculate the cross-correlation and find how the derivative of the cross-correlation function is linked to the Green's function.

We introduce the integral kernel of an operator P , denoted by $\llbracket P \rrbracket(x, y)$ by:

$$\forall u : X \longrightarrow \mathbb{R}^3, (Pu)(x) = \int_X \llbracket P \rrbracket(x, y) u(y) dy \quad (1.8)$$

This is the 'continuous matrix' of the operator P . It has to be linked to the case of a finite space where one can define the matrix (P_{ij}) of P and write the following formula:

$$\forall u : X \longrightarrow \mathbb{R}^3, (Pu)_i = \sum_j P_{ij} u_j \quad (1.9)$$

We first consider a medium without attenuation, *i.e.*, $a = 0$ in Eq. 1.7. Let us define the Green's function of L in the frequency domain, denoted $\hat{G}(\omega + i\varepsilon, \vec{r}, \vec{r}_s)$, with ε a small positive value, as the opposite of the integral kernel of $((\omega + i\varepsilon)^2 + L)^{-1}$. In other words, \hat{G} is the *resolvent* of L evaluated at point $(\omega + i\varepsilon)^2$. The $\hat{\cdot}$ denotes a function defined in the Fourier space. ε ensures that $((\omega + i\varepsilon)^2 + L)$ is invertible as L has real eigenvalues. We will show that this mathematical definition of \hat{G} is the same as the usual one, which is the causal solution of the wave equation (Eq. 1.7) when the source function f is a Dirac impulse in time and space $\delta(t, \vec{r} - \vec{r}_s)$. The Green's function $\hat{G}(\omega + i\varepsilon, \vec{r}, \vec{r}_s)$ admits a limit as $\varepsilon \rightarrow 0^+$, denoted $\hat{G}(\omega + i0, \vec{r}, \vec{r}_s)$, as a Schwartz distribution on the real axis. If L has a continuous spectrum, this limit is a smooth function (the 'limiting absorption principle'). \hat{G} can thus be written as:

$$\hat{G}(\omega + i0, \vec{r}, \vec{r}_s) = -\llbracket ((\omega + i0)^2 + L)^{-1} \rrbracket(\vec{r}, \vec{r}_s) \quad (1.10)$$

$$= -\int_X \llbracket ((\omega + i0)^2 + L)^{-1} \rrbracket(\vec{r}, \vec{r}') \delta(\vec{r}' - \vec{r}_s) d\vec{r}' \quad (1.11)$$

$$= -((\omega + i0)^2 + L)^{-1} \delta(\vec{r} - \vec{r}_s) \quad (1.12)$$

which yields:

$$-((\omega + i0)^2 + L) \hat{G}(\omega + i0, \vec{r}, \vec{r}_s) = \delta(\vec{r} - \vec{r}_s) \quad (1.13)$$

The inverse Fourier transform of this equation gives a relation that is the usual definition of G in the case of a medium without attenuation:

$$\frac{\partial^2 G}{\partial t^2}(t, \vec{r}, \vec{r}_s) - L G(t, \vec{r}, \vec{r}_s) = \delta(t) \delta(\vec{r} - \vec{r}_s) \quad (1.14)$$

G is thus the solution of Eq. 1.7 in the case of an impulsive source in time and space. One can compute the inverse Fourier transform of $\hat{G}(\omega + i\varepsilon, x, y)$ using residue calculus, and take the limit as ε goes to 0 to get

$$G(t, \vec{r}, \vec{r}_s) = Y(t) \left[\frac{\sin t \sqrt{-L}}{\sqrt{-L}} \right](\vec{r}, \vec{r}_s) \quad (1.15)$$

where Y is the Heaviside-step function, and where we denote $\sqrt{-L}$ the operator which eigenvalues are the images of the eigenvalues of L by the function $x \mapsto \sqrt{-x}$ (idem for the sinus function).

If we consider an attenuating medium, the Green's function $\hat{G}_a(\omega, \vec{r}, \vec{r}_s)$ is defined by the resolvent of L evaluated at point $\omega^2 + 2ia\omega$ instead of $(\omega + i0)^2$. It thus becomes

$$G_a(t, \vec{r}, \vec{r}_s) = Y(t) e^{-at} \left[\frac{\sin t \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right](\vec{r}, \vec{r}_s) \quad (1.16)$$

We now define the time domain cross-correlation between the displacement at two points A and B as:

$$C(\tau, \vec{r}_A, \vec{r}_B) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T u(t, \vec{r}_A) \overline{u(t + \tau, \vec{r}_B)} dt \quad (1.17)$$

where the bar denotes the conjugate. $u(t, \vec{r})$ can be expressed using the Green's function G_a (here attenuation is necessary to ensure convergence of the integral, see Roux et al. [2005b]) and the source function f as follows:

$$u(t, \vec{r}) = \int_0^\infty dt' \int_X G_a(t', \vec{r}, \vec{r}_s) f(t - t', \vec{r}_s) d\vec{r}_s \quad (1.18)$$

We assume that f is a white noise distributed everywhere in the medium X , acting at any time t . In the frequency domain, a white noise contains all the frequencies with a random phase. In the time domain, this is a random wavefield such that the position and the activation time of each source are uncorrelated. In this case, and considering a damping medium, we replace the large T limit in the correlation by an ensemble average. We then get the following explicit expression for the correlation between the wavefields recorded at A and B (see the last paragraph of this section for mathematical details):

$$C(\tau, \vec{r}_A, \vec{r}_B) = \frac{\sigma^2 e^{-a|\tau|}}{4a} \left[(-L)^{-1} \left(\cos \tau \sqrt{-L - a^2} + a \frac{\sin |\tau| \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right) \right] (\vec{r}_A, \vec{r}_B) \quad (1.19)$$

where σ is the variance of the noise wavefield.

The time derivative of this equation is expressed in terms of the Green's function using (1.16), giving the more familiar expression:

$$\frac{d}{d\tau} C(\tau, \vec{r}_A, \vec{r}_B) = \frac{-\sigma^2}{4a} (G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)) \quad (1.20)$$

This means that for any medium, the time-derivative of the cross-correlation computed between the wavefields recorded at two stations A and B is the Green's function of the medium, provided that the damping coefficient is small enough and that noise sources behave as white noise acting everywhere in the medium. This is the same hypothesis as stated in Roux et al. [2005b], Lobkis and Weaver [2001] and others, but L is now an arbitrary negative definite elliptic operator, and so the present result is more general.

The case of a scattering medium

The previous calculation was made using sources randomly located anywhere and randomly active at any time. This is a very strong hypothesis that is not valid in practical cases. Another demonstration of the link between cross-correlations and Green's functions can be made without any assumption about the noise sources location or their activation time. We only assume that there is equipartition at the boundaries of the region of interest, which means that each eigenmode is excited with the same level of energy.

A simple view of the relation between equipartition and correlation is given by the reconstruction of the Green's function of the homogeneous space using the azimuthal averaging of the correlation of plane waves, which are the eigenfunctions of the problem. Sánchez-Sesma and Campillo [2006] consider an isotropic distribution of P and S plane waves in an elastic medium. They found that the azimuthal average of the cross-correlation of motion between two points is proportional to the imaginary part of the exact Green's tensor between these points under the condition that the energy ratio S/P of the incident waves is the one predicted by equipartition. These results clearly show that equipartition is a necessary condition to retrieve the exact Green's function from correlations of the elastic field. In practice, one has to deal with complex media for which the eigenfunctions are unknown and therefore for which equipartition conditions cannot be explicitly specified in terms of local properties of the field.

Sánchez-Sesma et al. [2006] discussed a particular case. They consider the field in the vicinity of a cylindrical scatterer embedded in an homogeneous space and illuminated isotropically with incident P and S plane waves in the ratio of equipartition of the homogeneous space. Taking into account the scattered waves, they show that the azimuthal average of cross-correlations of motion between two points still yields the imaginary part of the exact Green's tensor of the heterogeneous medium, including the scattered waves, even at close distance from the scatterer. Is such a property still valid for any scattered or type of heterogeneity? What are the conditions required for the incident field? Weaver and Lobkis [2004] used an integral representation approach to study the problem of an heterogeneous region in an open medium. The essence of this property is expressed in the spectral theory of scattering that shows that the properties obtained in the simplest case of an homogeneous medium are formally valid in presence of heterogeneities. This is discussed in Colin de Verdière [2006a,b] as follows.

In a first step we will define the *spectral projector* and exhibit its expression using the cross-correlation function (Eq. 1.28). In a second step, will prove the relation between this projector and the Green's function, the so-called Stone formula (Eq. 1.35).

We denote $e_0(\vec{r}, \vec{k}) = e^{i\vec{k} \cdot \vec{r}}$ the plane waves that are the eigenmodes of the homogeneous infinite space. In the case of a complex medium, the scattering theory [Ramm, 1986, Reed and Simon, 1978] tells us that, if the medium is heterogeneous only in a finite region, the eigenmodes in the whole space can be written as:

$$e(\vec{r}, \vec{k}) = e_0(\vec{r}, \vec{k}) + e_s(\vec{r}, \vec{k}) \quad (1.21)$$

where e_s , the scattered waves, satisfies the so-called *Sommerfeld radiation* condition, which ensures that e_s will vanish when \vec{r} goes to infinity. This decomposition is still valid in the near field of the scatterers (*i.e.*, inside the heterogeneous region).

For $I \subset \mathbb{R}_+$, we define the spectral projector of L on I , denoted P_I , from its integral kernel $[[P_I]]$ at any points \vec{r}_1 and \vec{r}_2 of X by:

$$[[P_I]](\vec{r}_1, \vec{r}_2) = (2\pi)^{-d} \int_{\lambda_k \in I} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^d \vec{k}| \quad (1.22)$$

where d is the dimension of the space and λ_k denotes the eigenvalue associated with the eigenfunction $e(\vec{r}, \vec{k})$. This is the projector on the sub-eigenspaces of L which eigenvalues are in I . Again, what P_I represents can be easily understood in the case of a finite number N of eigenvalues $\{\lambda_n\}$, where we can write:

$$[[P_I]](\vec{r}_1, \vec{r}_2) = \sum_{\substack{n \in [1, N] \\ \lambda_n \in I}} e_i(\vec{r}_1) \overline{e_n(\vec{r}_2)} \quad (1.23)$$

In this case, I represents a subset of values of $n \in [1, N]$ that are preserved, all the other being removed by the projector P_I . For example, if $u(\vec{r}) = \sum_{n=1}^N u_n e_n(\vec{r})$, we have:

$$(P_I u)(\vec{r}) = \sum_{\substack{n \in [1, N] \\ \lambda_n \in I}} u_n e_n(\vec{r}) \quad (1.24)$$

We now will demonstrate that the derivative of the spectral projector on an interval around a value $\omega^2 \in \mathbb{R}_+$ is linked to the cross-correlation function at the corresponding pulsation ω . We

thus consider an interval $I = [\omega_-^2, \omega_+^2]$ around ω^2 . The integral over $\lambda_k = c^2 |\vec{k}|^2 \in I$ in Eq. 1.22 defines a volume of integration of dimension d that can be decomposed into two integrals, one over a volume of dimension $d - 1$ defined by $c^2 |\vec{k}|^2 = \omega^2$ and the other over $|\vec{k}|$:

$$\llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = (2\pi)^{-d} \int_{c^2 |\vec{k}|^2 \in I} \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^{d-1}\sigma| |\vec{k}|^{d-1} d|\vec{k}| \quad (1.25)$$

where $|d^{d-1}\sigma|$ is the usual measure of the unit $(d - 1)$ -dimensional sphere. In the case $d = 3$, $d^2\sigma$ is the infinitesimal solid angle. Taking the derivative with respect to ω_+ in this equation gives:

$$\frac{d}{d\omega_+} \llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = \frac{(2\pi)^{-d}}{c} \left(\frac{\omega}{c}\right)^{d-1} \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^{d-1}\sigma| \quad (1.26)$$

In this formula, one can recognize the cross-correlation of random scattered waves of frequency ω recorded at points \vec{r}_1 and \vec{r}_2 that can be written as:

$$\hat{C}(\omega, \vec{r}_1, r_2) = \frac{1}{\sigma_{d-1}} \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d\sigma| \quad (1.27)$$

where σ_{d-1} denote the total volume of the unit sphere in \mathbb{R}^{d-1} : $\sigma_0 = 2$, $\sigma_1 = 2\pi$, $\sigma_2 = 4\pi$, ...

Using the two previous equations, we find

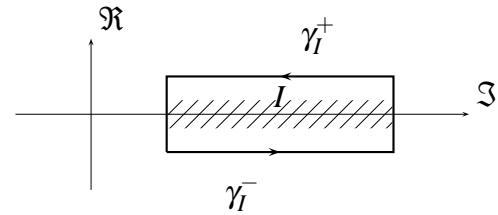
$$\frac{d}{d\omega} \llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = \frac{\sigma_{d-1}}{(2\pi)^d} \frac{1}{c} \left(\frac{\omega}{c}\right)^{d-1} C_\omega(\vec{r}_1, \vec{r}_2) \quad (1.28)$$

The projector P_I defined previously can also be written using the resolvent of the operator L using the Cauchy formula:

$$P_I = \frac{1}{2i\pi} \int_{\gamma} (L + \lambda)^{-1} d\lambda \quad (1.29)$$

where a γ is a contour in the complex plane which restriction to the real axis is I .

Figure 1.5: γ can be split into two contours γ_I^+ and γ_I^- which are complex conjugate and followed in opposite direction.



This contour can be split into two contours defined by $\gamma_I^+ = \{\lambda \in \gamma \mid \Im(\lambda) \geq 0\}$ and $\gamma_I^- = \{\lambda \in \gamma \mid \Im(\lambda) < 0\}$ (\Im denotes the imaginary part) as seen in Fig. 1.5. As γ_I^+ and γ_I^- are followed in opposite directions and as they are complex conjugates, we get:

$$P_I = \frac{1}{2i\pi} \int_{\gamma_I^+} [(L + \lambda)^{-1} - (L + \bar{\lambda})^{-1}] d\lambda \quad (1.30)$$

$$= \frac{1}{\pi} \int_{\gamma_I^+} \Im(L + \lambda)^{-1} d\lambda \quad (1.31)$$

and then, taking the integral kernel of this expression,

$$\llbracket P_I \rrbracket = \frac{1}{\pi} \int_{\gamma_\varepsilon^+} \Im \llbracket (L + \lambda)^{-1} \rrbracket d\lambda \quad (1.32)$$

$$= -\frac{1}{\pi} \int_{\omega_-}^{\omega_+} \Im \llbracket (L + (\omega + i0)^2)^{-1} \rrbracket 2\omega d\omega \quad (1.33)$$

which leads to the Stone formula, using the definition of \hat{G}

$$\llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = -\frac{2}{\pi} \int_{\omega_-}^{\omega_+} \omega \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) d\omega \quad (1.34)$$

This formula gives, by taking the derivative with respect to ω_+ ,

$$\frac{d}{d\omega_+} \llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = -\frac{2\omega}{\pi} \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) \quad (1.35)$$

The combination of Eq. 1.28 and Eq. 1.35 finally gives:

$$\hat{C}(\omega, \vec{r}_1, \vec{r}_2) = -\frac{2^{d+1} \pi^{d-1}}{\sigma_{d-1}} \frac{c^d}{\omega^{d-2}} \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) \quad (1.36)$$

This gives a generalization of Eq. 1.20 in the case of observation in a region without local sources, and requires no hypothesis about attenuation. It shows that the equipartition at boundaries of the region of interest is sufficient to obtain the Green's function from cross-correlation, whatever is the wavefield inside the medium. This equation, established in the scalar case, can be extended to the elastic case using the same calculation. The cross-correlation function becomes a tensor, as well as the Green's function. Particular attention needs to be paid to velocities and dispersion relations as they depend on the type of waves.

The rate of convergence towards the Green's function

The question arises as to how much averaging is in principle necessary after which the Green's function is retrieved, and before which the cross-correlation remains dominated by noise. In the case that the field is due to uniformly distributed random sources, or in the case that the field is due to equipartitioned incident waves, it is possible to make theoretical estimates [Larose et al., 2004, Snieder, 2004, Sabra et al., 2005c, Weaver and Lobkis, 2005a]. All these authors have concluded, not surprisingly, that the convergence proceeds like the square root of the amount of data used in the cross-correlation. Signal-to-Noise Ratio (SNR), *i.e.*, Green's function amplitude over residual fluctuations, is proportional to this square root. Quantitative estimates of the quality of the convergence are more challenging. Weaver and Lobkis [2005a] calculated the residual error in a scalar wave cross-correlation, and found it to be proportional to the energy in the diffuse field times the bandwidth times the fourth power of central frequency. A similar calculation for closed systems was confirmed in laboratory measurements [Weaver and Lobkis, 2005b]. The residual error was compared to the amplitude of a ray arrival expected in the converged cross-correlation. Each ray arrival amplitude A depends on 1) the geometrical spreading of the Green's function, and 2) the spatial extension of the noise sources that coherently contribute to the Green's function reconstruction.

This zone is characterized by a directivity angle $\delta\theta = \sqrt{\frac{c}{r\omega}}$, where r is the source-receiver distance. The ray arrival was shown to be apparent in the cross-correlation if $\delta t \delta\omega \gg A^{d-1}$, where δt is the amount of data record employed (this is often months in seismic applications), $\delta\omega$ is the bandwidth of interest (often around 1 Hz or less in seismic applications), and $A = r\omega/c$, (the source-receiver distance r times the wavenumber $k = 2\pi/\lambda$). The power is equal to one less than the dimension d of the propagation; thus $d - 1 = 1$ for Rayleigh waves. Propagation between distant source-receiver pairs, and propagation in three dimensions, are especially challenging to resolve, largely due to the weakness of such ray arrivals.

Appendix: Detailed calculation of the cross-correlation function (Eq. 1.19)

We start from the definition of the cross-correlation function between two points A and B (Eq. 1.17) in which we express the wavefield u using the Green's function G :

$$C(\tau, \vec{r}_A, \vec{r}_B) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T u(t, \vec{r}_A) \overline{u(t + \tau, \vec{r}_B)} dt \quad (1.37)$$

$$\begin{aligned} &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T dt \int_0^\infty ds \int_X d\vec{r}_s G_a(s, \vec{r}_A, \vec{r}_s) f(t - s, \vec{r}_s) \\ &\quad \times \int_0^\infty ds' \int_X d\vec{r}_{s'} \overline{G_a(s', \vec{r}_B, \vec{r}_{s'}) f(t + \tau - s', \vec{r}_{s'})} \end{aligned} \quad (1.38)$$

The large T limit can be replaced by an ensemble average, that gives the mathematical expectation denoted \mathbb{E} . As f is a white noise, we have:

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(t - s, \vec{r}_s) f(t + \tau - s', \vec{r}_{s'}) dt = \mathbb{E}[f(t - s, \vec{r}_s) f(t + \tau - s', \vec{r}_{s'})] \quad (1.39)$$

$$= \sigma^2 \delta(\tau + s - s') \delta(\vec{r}_s - \vec{r}_{s'}) \quad (1.40)$$

where σ is the variance of the white noise. This property simplifies the previous equation, and we get:

$$C(\tau, \vec{r}_A, \vec{r}_B) = \sigma^2 \int_0^\infty ds \int_X d\vec{r}_s G_a(s, \vec{r}_A, \vec{r}_s) \overline{G_a(s + \tau, \vec{r}_B, \vec{r}_s)} \quad (1.41)$$

Using the expression of the Green's function (Eq. 1.16):

$$\begin{aligned} C(\tau, \vec{r}_A, \vec{r}_B) &= \sigma^2 \int_0^\infty ds \int_X d\vec{r}_s Y(s) Y(s + \tau) e^{-as} e^{-a(s+\tau)} \\ &\quad \left[\frac{\sin s \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_A, \vec{r}_s) \overline{\left[\frac{\sin(s + \tau) \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_B, \vec{r}_s)} \end{aligned} \quad (1.42)$$

We use two properties of the integral kernel:

$$\overline{[P]}(x, y) = [\overline{P}](x, y) = [P](y, x) \quad (1.43)$$

$$\int_X [P_1](x, z) [P_2](z, y) dz = [P_1 \cdot P_2](x, y) \quad (1.44)$$

to get a new formula for the cross-correlation function:

$$\begin{aligned} C(\tau, \vec{r}_A, \vec{r}_B) &= \sigma^2 \int_0^\infty ds Y(s + \tau) e^{-a(2s+\tau)} \\ &\quad \left[\frac{\sin s \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \frac{\sin(s + \tau) \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_A, \vec{r}_B) \end{aligned} \quad (1.45)$$

Using $\sin \alpha \sin \beta = 1/2 (\cos(\alpha - \beta) - \cos(\alpha + \beta))$ and computing the integral over ds , we get Eq. 1.19:

$$C(\tau, \vec{r}_A, \vec{r}_B) = \frac{\sigma^2 e^{-a|\tau|}}{4a} \left[(-L)^{-1} \left(\cos \tau \sqrt{-L - a^2} + a \frac{\sin |\tau| \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right) \right] (\vec{r}_A, \vec{r}_B) \quad (1.19)$$

Bibliographie chapitre 1

- K. Aki. Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, 35:415–456, 1957.
- K. Aki. Analysis of the seismic coda of local earthquakes as scattered waves. *Journal of Geophysical Research*, 74:615–618, 1969.
- K. Aki and B. Chouet. Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(9):3322–3342, Aug 1975.
- G. D. Bensen, M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260, Jun 2007. doi: 10.1111/j.1365-246X.2007.03374.x.
- M. Campillo. Phase and correlation in ‘random’ seismic fields and the reconstruction of the Green function. *Pure and Applied Geophysics*, 163(2-3):475–502, Mar 2006. doi: 10.1007/s00024-005-0032-8.
- M. Campillo and A. Paul. Long-range correlations in the diffuse seismic coda. *Science*, 299: 547–549, Jan 2003. doi: 10.1126/science.1078551.
- M. Campillo, L. Margerin, and N. M. Shapiro. *Seismic wave diffusion in the Earth lithosphere*. NATO ASI Series, Kluwer, 1999.
- D. Cassereau and M. Fink. Time-reversal of ultrasonic fields – Part III: Theory of the closed time-reversal cavity. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(5): 579–592, Sep 1992. doi: 10.1109/58.156176.
- J. F. Claerbout. Synthesis of a layered medium from its acoustics transmission response. *Geophysics*, 33:264–269, 1968.
- Y. Colin de Verdière. Mathematical models for passive imaging I: general background, 2006a. URL <http://fr.arxiv.org/abs/math-ph/0610043/>.
- Y. Colin de Verdière. Mathematical models for passive imaging II: effective hamiltonians associated to surface waves, 2006b. URL <http://fr.arxiv.org/abs/math-ph/0610044/>.
- Y. Colin de Verdière. Semi-classical analysis and passive imaging, 2008. URL <http://fr.arxiv.org/abs/0803.2436>. Submitted to Nonlinearity.
- M. L. Cowan, J. H. Page, and D. A. Weitz. Velocity fluctuations in fluidized suspensions probed by ultrasonic correlation spectroscopy. *Physical Review Letters*, 85(2):453–456, Jul 2000.
- A. Derode, E. Larose, M. Campillo, and M. Fink. How to estimate the Green’s function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Applied Physics Letters*, 83(15):3054–3056, Oct 2003a. doi: 10.1063/1.1617373.

- A. Derode, E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink. Recovering the Green's function from field-field correlations in an open scattering medium (L). *The Journal of the Acoustical Society of America*, 113(6):2973–2976, 2003b. doi: 10.1121/1.1570436.
- T. L. Duvall, S. M. Jefferies, J. W. Harvey, and M. A. Pomerantz. Time distance helioseismology. *Nature*, 362:430–432, Apr 1993. doi: 10.1038/362430a0.
- M. Fink. Time reversal of ultrasonic fields – Part I: Basic principles. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(5):555–566, Sep 1992.
- M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L. Thomas, and F. Wu. Time-reversed acoustics. *Reports on Progress in Physics*, 63:1933–1995, 2000. doi: 10.1088/0034-4885/63/12/202.
- P. M. Gilles, T. L. Duvall, P. H. Scherrer, and R. S. Bogart. A subsurface flow of material from the Sun's equator to its poles. *Nature*, 390:52–54, Nov 1997. doi: 10.1038/36294.
- O. A. Godin. Emergence of the acoustic Green's function from thermal noise. *The Journal of the Acoustical Society of America*, 121(2):EL96–EL102, Feb 2007. doi: 10.1121/1.2430764.
- P. Gouédard, L. Stehly, F. Brenguier, M. Campillo, Y. Colin de Verdière, E. Larose, L. Margerin, P. Roux, F. J. Sánchez-Sesma, N. M. Shapiro, and R. L. Weaver. Cross-correlation of random fields: mathematical approach and applications. *Geophysical Prospecting*, 56(3):375–393, May 2008. doi: 10.1111/j.1365-2478.2007.00684.x.
- R. Hennino, N. Tégourès, N. M. Shapiro, L. Margerin, M. Campillo, B. A. van Tiggelen, and R. L. Weaver. Observation of equipartition of seismic waves. *Physical Review Letters*, 86(15):3447–3450, Apr 2001. doi: 10.1103/PhysRevLett.86.3447.
- R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255–284, 1966. doi: 10.1088/0034-4885/29/1/306.
- E. Larose, A. Derode, M. Campillo, and M. Fink. Imaging from one-bit correlation of wide-band diffuse wavefield. *Journal of Applied Physics*, 95(12):8393–8399, 2004. doi: 10.1063/1.1739529.
- E. Larose, A. Khan, Y. Nakamura, and M. Campillo. Lunar subsurface investigated from correlation of seismic noise. *Geophysical Research Letters*, 32:L16201, 2005. doi: 10.1029/2005GL023518.
- E. Larose, L. Margerin, A. Derode, B. A. van Tiggelen, M. Campillo, N. M. Shapiro, A. Paul, L. Stehly, and M. Tanter. Correlation of random wavefields: an interdisciplinary review. *Geophysics*, 71(4):SI11–SI21, Jul-Aug 2006. doi: 10.1190/1.2213356.
- O. I. Lobkis and R. L. Weaver. On the emergence of the Green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6):3011–3017, Dec 2001. doi: 10.1121/1.1417528.
- L. Margerin, M. Campillo, N. M. Shapiro, and B. A. van Tiggelen. Residence time of diffuse waves in the crust as a physical interpretation of coda Q : application to seismograms recorded in Mexico. *Geophysical Journal International*, 138(2):343–352, 1999.

- L. Margerin, M. Campillo, and B. A. van Tiggelen. Monte carlo simulation of multiple scattering of elastic waves. *Journal of Geophysical Research*, 105(B4):7873–7892, Apr 2000.
- A. Paul, M. Campillo, L. Margerin, E. Larose, and A. Derode. Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves. *Journal of Geophysical Research (Solid Earth)*, 110:B08302, Aug 2005. doi: 10.1029/2004JB003521.
- G. Poupinet, W. L. Ellsworth, and J. Frechet. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California. *Journal of Geophysical Research*, 89(B7):5719–5731, 1984.
- A. G. Ramm. *Scattering by obstacles*. Kluwer Academic Publisher, 1986.
- M. Reed and B. Simon. *Methods of modern mathematical physics I, II, III, IV*. Academic Press, 1978.
- P. M. Roberts, W. S. Phillips, and M. C. Fehler. Development of the active doublet method for measuring small velocity and attenuation changes in solids. *The Journal of the Acoustical Society of America*, 91(6):3291–3302, Jun 1992.
- P. Roux and W. A. Kuperman. Extracting coherent wavefronts from acoustic ambient noise in the ocean. *The Journal of the Acoustical Society of America*, 116(4):1995–2003, Oct 2004. doi: 10.1121/1.1797754.
- P. Roux, K. G. Sabra, P. Gerstoft, and W. A. Kuperman. P-waves from cross-correlation of seismic noise. *Geophysical Research Letters*, 32:L19303, 2005a. doi: 10.1029/2005GL023803.
- P. Roux, K. G. Sabra, W. A. Kuperman, and A. Roux. Ambient noise cross correlation in free space: theoretical approach. *The Journal of the Acoustical Society of America*, 117(1):79–84, Jan 2005b. doi: 10.1121/1.1830673.
- L. Ryzhik, G. Papanicolaou, and J. B. Keller. Transport equations for elastic and other waves in random media. *Wave motion*, 24:327–370, 1996.
- K. G. Sabra, P. Gerstoft, P. Roux, and W. A. Kuperman. Surface wave tomography from microseisms in southern California. *Geophysical Research Letters*, 32:L14311, Jul 2005a. doi: 10.1029/2005GL023155.
- K. G. Sabra, P. Roux, and W. A. Kuperman. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide. *The Journal of the Acoustical Society of America*, 117(1):164–174, Jan 2005b. doi: 10.1121/1.1835507.
- K. G. Sabra, P. Roux, and W. A. Kuperman. Emergence rate of the time-domain Green’s function from the ambient noise cross-correlation. *The Journal of the Acoustical Society of America*, 118(6):3524–3531, Dec 2005c. doi: 10.1121/1.2109059.
- F. J. Sánchez-Sesma and M. Campillo. Retrieval of the Green function from cross correlation: the canonical elastic problem. *Bulletin of the Seismological Society of America*, 96(3):1182–1191, Jun 2006. doi: 10.1785/0120050181.

- F. J. Sánchez-Sesma, J. A. Pérez-Ruiz, M. Campillo, and F. Luzón. Elastodynamic 2-D Green function retrieval from cross-correlation: canonical inclusion problem. *Geophysical Research Letters*, 33:L13305, 2006. doi: 10.1029/2006GL026454.
- F. J. Sánchez-Sesma, J. A. Pérez-Ruiz, F. Luzón, M. Campillo, and A. Rodríguez-Castellanos. Diffuse fields in dynamic elasticity. *Wave Motion*, 45(5):641–654, Apr 2008. doi: 10.1016/j.wavemoti.2007.07.005.
- N. M. Shapiro and M. Campillo. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31:L07614, 2004. doi: 10.1029/2004GL019491.
- N. M. Shapiro, M. Campillo, L. Margerin, S. K. Singh, V. Kostoglodov, and J. Pacheco. The energy partitioning and the diffusive character of the seismic coda. *Bulletin of the Seismological Society of America*, 90(3):655–665, Jun 2000. doi: 10.1785/0119990021.
- N. M. Shapiro, M. Campillo, L. Stehly, and M. H. Ritzwoller. High-resolution surface wave tomography from ambient seismic noise. *Science*, 307:1615–1618, Mar 2005. doi: 10.1126/science.1108339.
- R. Snieder. Coda wave interferometry and the equilibration of energy in elastic media. *Physical Review E*, 66:046615, 2002. doi: 10.1103/PhysRevE.66.046615.
- R. Snieder. Extracting the Green’s function from the correlation of coda waves: a derivation based on stationary phase. *Physical Review E*, 69(4):046610, Apr 2004. doi: 10.1103/PhysRevE.69.046610.
- R. Snieder. The theory of coda wave interferometry. *Pure and Applied Geophysics*, 163(2):455–473, Mar 2006. doi: 10.1007/s00024-005-0026-6.
- L. Stehly, M. Campillo, and N. M. Shapiro. A study of the seismic noise from its long range correlation properties. *Journal of Geophysical Research*, 111:B10306, 2006. doi: 10.1029/2005JB004237.
- L. Stehly, M. Campillo, and N. M. Shapiro. Travel time measurements from noise correlation: stability and detection of instrumental time-shifts. *Geophysical Journal International*, 171(1): 223–230, Oct 2007. doi: 10.1111/j.1365-246X.2007.03492.x.
- A. Tourin, A. Derode, and M. Fink. Sensitivity to perturbations of a time-reversed acoustic wave in a multiple scattering medium. *Physical Review Letters*, 87(27):274301, 2001. doi: 10.1103/PhysRevLett.87.274301.
- B. A. van Tiggelen. Green function retrieval and time reversal in a disordered world. *Physical Review Letters*, 91(24):243904, Dec 2003. doi: 10.1103/PhysRevLett.91.243904.
- K. Wapenaar. Retrieving the elastodynamic Green’s Function of an arbitrary inhomogeneous medium by cross-correlation. *Physical Review Letters*, 93(25):254301, Dec 2004. doi: 10.1103/PhysRevLett.93.254301.

- K. Wapenaar. Nonreciprocal Green's function retrieval by cross correlation. *The Journal of the Acoustical Society of America*, 120(1):EL7–EL13, Jul 2006. doi: 10.1121/1.2208153.
- R. L. Weaver. Information from seismic noise. *Science*, 307(5715):1568–1569, Mar 2005. doi: 10.1126/science.1109834.
- R. L. Weaver and O. I. Lobkis. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies. *Physical Review Letters*, 87(13):134301, Sep 2001. doi: 10.1103/PhysRevLett.87.134301.
- R. L. Weaver and O. I. Lobkis. Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal phonons. *The Journal of the Acoustical Society of America*, 113(5):2611–2621, 2003. doi: 10.1121/1.1564017.
- R. L. Weaver and O. I. Lobkis. Diffuse fields in open systems and the emergence of the Green's function. *The Journal of the Acoustical Society of America*, 116(5):2731–2734, Nov 2004. doi: 10.1121/1.1810232.
- R. L. Weaver and O. I. Lobkis. Fluctuations in diffuse field-field correlations and the emergence of the Green's function in open systems. *The Journal of the Acoustical Society of America*, 117(6):3432–3439, 2005a. doi: 10.1121/1.1898683.
- R. L. Weaver and O. I. Lobkis. The mean and variance of diffuse field correlations in finite bodies. *The Journal of the Acoustical Society of America*, 118(6):3447–3456, 2005b. doi: 10.1121/1.2109307.

Convergence vers la fonction de Green

Sommaire

2.1	Description du jeu de données Qarn Alam	44
2.2	Convergence of the two-point correlation function toward the Green's function in the context of a seismic prospecting dataset	44
2.2.1	Abstract	45
2.2.2	Introduction	45
2.2.3	The seismic network	47
2.2.4	Correlation function vs source distribution	47
2.2.5	Amplitude of the crosscorrelation function	52
2.2.6	Conclusion	54
2.2.7	Acknowledgments	54
2.3	Imagerie passive d'un milieu homogène	54
2.3.1	Analyse fréquence/nombre d'onde (FK)	54
2.3.2	Application de l'analyse FK aux données Qarn Alam	57
2.4	Tomographie passive d'ondes de surface	60
2.4.1	Algorithme d'inversion	62
2.4.2	Inversion de la zone de Qarn Alam	64
2.4.3	Conclusion et perspectives sur l'imagerie passive	70
2.5	Caractérisation de la diffusion : mesure du libre parcours moyen élastique	71
	Bibliographie chapitre 2	76

Ce chapitre est consacré à la convergence de la fonction de corrélation vers la fonction de Green, c'est-à-dire à l'étude des conditions nécessaires à la reconstruction de cette fonction. Les paramètres abordés ici seront la distribution des sources de bruit et l'importance de la diffusion dans le milieu de propagation. Nous avons vu que les hypothèses de base du théorème de corrélation sont un champ diffus équipartitionné et isotrope. Ces hypothèses sont rarement vérifiées dans la pratique et il est donc indispensable de comprendre les conséquences des écarts au cas idéal.

Le jeu de données utilisé pour cette étude est un jeu de données actives obtenu dans le cadre d'une collaboration avec Shell R&D. Sa particularité est la très grande densité de sources et de ré-

cepteurs déployés, comme décrit dans le paragraphe 2.1. J'ai en particulier profité de cette propriété en sélectionnant différentes distributions de sources afin d'analyser, entre autre choses, l'influence d'un défaut d'isotropie du champ d'onde sur le résultat de la corrélation. J'ai donc appliqué à ces données actives les mêmes traitements que ceux utilisés pour les enregistrements passifs. Plus que la simple convergence vers la fonction de Green, je me suis intéressé à l'application de méthodes d'imagerie géophysique, et à la comparaison entre les résultats « passifs » (*i.e.*, obtenus avec des données actives traitées comme l'auraient été des données passives) et les résultats actifs. En effet ce jeu de données inclus des sources situées à proximité de chaque récepteur, ce qui permet d'avoir des réponses impulsionnelles de référence afin de vérifier la convergence de la fonction de corrélation vers la fonction de Green, mais aussi d'appliquer les mêmes algorithmes d'imagerie aux données actives et passives pour le même milieu. Les dégradations de l'image finale obtenue en fonction de l'écart aux hypothèses de base du théorème de corrélation seront caractérisées.

2.1 Description du jeu de données Qarn Alam

Une étude haute résolution a été réalisée par *Petroleum Development Oman* (PDO) dans la région de Qarn Alam, au nord de l'Oman. Les données ont été acquises dans une zone d'un kilomètre carré, couverte par 1600 géophones placés aux nœuds d'une grille de 25 m d'espacement. Durant cette campagne d'acquisition active, 1600 sources vibratoires ont été placées tour à tour sur les nœuds d'une grille similaire, décalée d'une demi cellule (12,5 m) dans les deux directions par rapport à la grille des récepteurs. Les sources utilisées sont des camions vibrateurs travaillant dans la bande de fréquences 8-120 Hz, mais les traitements présentés dans ce manuscrit ont été réalisés dans la bande 10-25 Hz (sauf mention contraire). Le jeu de données complet est ainsi constitué de 1600×1600 signaux temporels, et il constitue une mesure exhaustive des fonctions de transfert dans la zone étudiée. Le milieu est karstique et donc diffusant. Les enregistrements ont une durée de 4 secondes, et présentent de la coda aux temps longs, comme le montre la figure 2.1a.

Ce jeu de données nous permet de tester le fonctionnement des méthodes de corrélation en contrôlant la distribution de sources utilisée, d'une part, mais aussi l'importance de la diffusion à travers la sélection de fenêtres temporelles plus ou moins tardives dans les enregistrements. Ces données actives seront donc utilisées comme l'auraient été des enregistrements de bruit, avec l'avantage de mieux contrôler les propriétés du champ d'onde.

2.2 Convergence of the two-point correlation function toward the Green's function in the context of a seismic prospecting dataset

Voici un article publié dans la revue *Geophysics* [Gouédard, Roux, Campillo, & Verdel, 2008a]. Dans cet article, nous avons voulu étudier l'influence de la distribution de sources et de la diffusion dans la reconstruction des fonction de Green par corrélation. Nous avons ainsi montré que, lorsque l'on utilise des ondes directes, la distribution de sources contrôle le résultat de la corrélation, et on ne retrouve la fonction de Green que lorsque les sources entourent complètement la

paire de capteurs utilisés, ou bien lorsqu'elles sont situées uniquement dans l'alignement de cette dernière. Inversement, lorsque l'on utilise les ondes diffusées de la coda, la distribution de sources utilisée n'a que peu d'influence puisque la diffusion a rendu le champ isotrope (pour des temps d'enregistrement supérieurs au temps libre moyen de transport). Nous montrons également que non seulement la phase, mais aussi l'amplitude relative de la fonction de Green peut être reconstruite pour des paires de récepteurs plus ou moins éloignés. Ce dernier résultat n'est valable que lorsque l'on enregistre un champ d'onde isotrope, obtenu grâce à une distribution homogène de sources ou à la diffusion. Lorsque les sources sont uniquement situées dans l'alignement des récepteurs, le processus de corrélation ne peut reconstruire qu'une onde plane se propageant dans la direction définie par les récepteurs, et la décroissance d'amplitude avec la distance n'inclut plus de dispersion géométrique.

► 2.2.1 Abstract

When considering direct waves in the correlation process, the Green's function is reconstructed when using an even distribution of seismic sources or when the source distribution is restricted to the direction close to the alignment of the sensors. On the other hand, when considering records of coda waves, the convergence is achieved for any source distribution, as expected theoretically. We extract the expected amplitude decay along a seismic profile from the correlation functions when an even distribution of sources is considered or when the time window includes scattered waves.

► 2.2.2 Introduction

The Green's function of a medium between points A and B represents the record we would get at A if an impulse source were applied at B. In the case of random isotropic noise sources, the cross-correlation of wavefields recorded at two points converges to the Green's function of the medium, including all reflection, scattering, and propagation modes. Various experimental, numerical, and theoretical approaches have been developed to demonstrate this result and to define more precisely under which assumption it is valid.

Aki [1957] proposed to use seismic noise to retrieve, frequency by frequency, the dispersion properties of surface waves in the subsoil. Later, helioseismology was the first field to use ambient noise crosscorrelation in the time domain for imaging. The recordings of random motion of the solar surface were processed to retrieve time-distance information on the solar subsurface [Duvall et al., 1993, Gilles et al., 1997, Rickett & Claerbout, 1999]. The same idea developed as Claerbout [1968] proposed using daylight imaging in geophysical prospecting.

More recently, a seminal paper by Weaver & Lobkis [2001] shows how diffuse thermal noise recorded and crosscorrelated at two ultrasonic transducers, working at megahertz frequencies and fastened to one face of a duralumin cube 10 cm on a side, provides the complete time-domain Green's function between these two points. This spectacular result has been the start of a complete revival of the subject because experimental evidence of ambient noise crosscorrelations have been demonstrated in acoustics and elastic plates [Lobkis & Weaver, 2001, Weaver & Lobkis, 2001, Larose et al., 2004, 2007, Sabra et al., 2008]; seismology [Paul & Campillo, 2001, Campillo &

Paul, 2003, Schuster et al., 2004, Shapiro & Campillo, 2004], where passive seismic imaging was performed in California from seismic noise at stations separated by hundreds to thousands of kilometers [Shapiro et al., 2005, Sabra et al., 2005a]; seismic exploration, where seismic body-wave reflections have been retrieved from noise measurements by an array of three-component geophones laid out in a desert area [Draganov et al., 2007]; and oceanography, where both direct and reflected wavefronts have been retrieved from ambient noise crosscorrelation in shallow underwater acoustics [Roux & Kuperman, 2004, Sabra et al., 2005b].

The first Weaver & Lobkis [2001] experimental result was interpreted theoretically by invoking equipartitioning of the modes excited in the aluminum sample. A spectral representation [Lobkis & Weaver, 2001], the fluctuation-dissipation approach [Weaver & Lobkis, 2001, 2003, van Tiggelen, 2003, Godin, 2007], or a correlation representation theorem [e.g., Wapenaar, 2004] yields rigorous theoretical approaches to interpreting experimental results in opened complex media. Cases of non-reciprocal (e.g., in the presence of a flow) or inelastic media have also been investigated [Wapenaar, 2006, Godin, 2007, Colin de Verdière, 2006a,b].

For the problem of elastic waves, it has been shown theoretically that the convergence of noise correlation to the Green's function is bonded by the equipartition condition of the different components of the elastic field [Sánchez-Sesma & Campillo, 2006, Sánchez-Sesma et al., 2008]. In other words, the emergence of the Green's function is effective after a sufficient self-averaging process that is provided by random spatial distribution of the noise sources when considering long time series as well as scattering [Campillo, 2006, Larose et al., 2006, Gouédard et al., 2008b].

In some theoretical approaches, Huygens's principle is used to replace the 3D-distributed ambient noise [Roux et al., 2005] by a distribution of sources on a surface surrounding the medium [Wapenaar et al., 2005]. As a matter of fact, this result has been demonstrated where experimental randomization is not produced by the distribution of sources but is provided by multiple scattering in heterogeneous media [Campillo & Paul, 2003]. By summing the contributions of a spatially uniform distribution of sources to the correlation, it has been shown numerically that the correlation contains the causal and acausal Green's function of the medium [Wapenaar, 2004]. Derode et al. [2003a,b] propose to interpret the Green's function reconstruction in terms of a time-reversal analogy and show that correlation of multiply scattered waves could be used for passive imaging in acoustics. Regarding the physics, the convergence of the correlation function toward the Green's function in an unbounded medium finds a natural interpretation through the stationary-phase theorem [Snieder, 2004].

Following these numerical results, we experimentally investigate the link between the spatial distribution of sources and the two-point correlation function when direct waves and multiply scattered waves are supported by the propagation medium. Our objective is to determine the source distribution to retrieve the Green's function for direct and diffuse wavefields. We take advantage of an extensive seismic-prospecting experimental data set where 1600 seismic sources and receivers were deployed on a 25×25 -m cell grid, as described in the following section. In the next section, we show the influence of scattering and source distribution on the convergence toward the Green's function. The last section is devoted to reconstructing relative amplitudes of the Green's function with correlation.

► 2.2.3 The seismic network

Petroleum Development Oman PDO recorded a high-resolution survey in northern Oman. The data were obtained in a 1×1 -km area covered with 1600 geophones located on a 25×25 -m cell grid. During this active-seismic survey, vibrators were located on a similar grid shifted with respect to the receiver grid by half a cell (12.5 m) in both directions. The complete data set was then made of 1600×1600 source-receiver time-domain signals constituting an exhaustive measurement of the transfer function of the half-space medium over a 1-km^2 area.

We first selected sources located on a 600-m-internal-diameter, 70-m-wide ring centered on a receiver (Figure 2.1a). For a given recording time at this receiver, waves from each source were

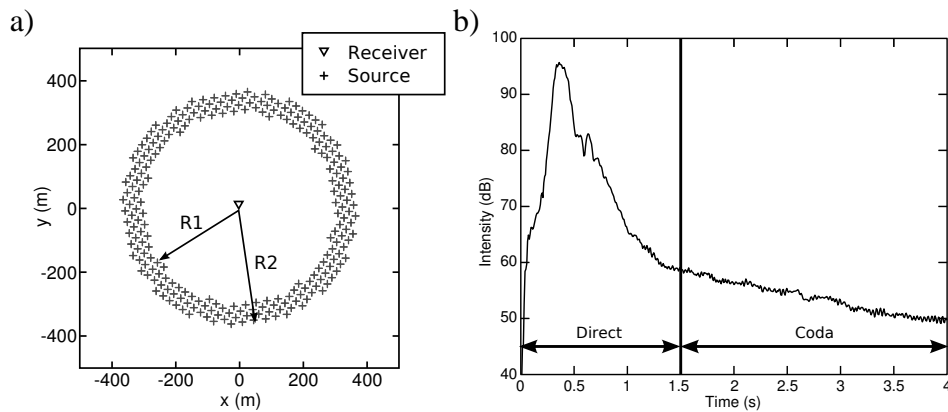


Figure 2.1: **a.** Experimental geometry when considering sources surrounding one receiver. Here, 240 sources are selected, located between two circles of radius $R1 = 300$ m and $R2 = 370$ m, centered on the receiver. **b.** Intensity received at the receiver averaged over the sources selected in (a), in decibel scale.

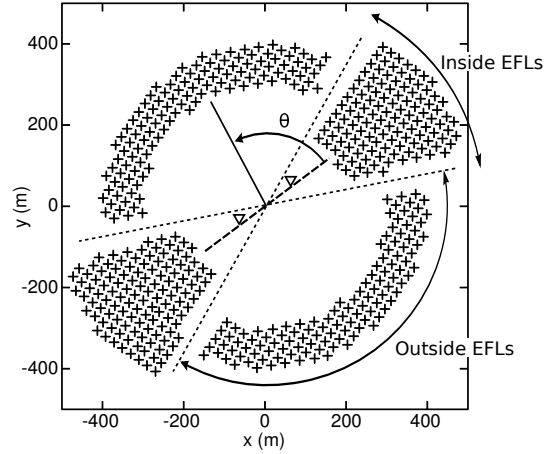
in the same scattering regime because the source-receiver distance was about the same for each source. Figure 2.1b shows the intensity received at the receiver averaged over all sources in the ring. The shallow subsurface appears to be heterogeneous, and the 4-s records include seismic coda. Direct surface waves arrived in about 0.4 s, which corresponds to a 1000-m/s average speed in the 15–45-Hz frequency bandwidth. At times greater than 1.5 s, the energy decay was linear in decibel scale. This indicated the multiple scattering regime of coda waves. An experimental study had shown that the mean free path in this area was around 1 km in this frequency range. Selecting a time window in the recorded signals was a way to choose the number of scattered waves in the correlation process. Correlation is equivalent to seismic interferometry [Schuster et al., 2004] when it is performed with direct waves and correlated sources; correlation resembles coda-wave interferometry when the diffuse field from uncorrelated sources is involved. In the following, we focus on surface waves, which are the predominant direct arrivals on the seismograms and play a major part in the coda [Hennino et al., 2001, Campman et al., 2005, Herman & Perkins, 2006].

► 2.2.4 Correlation function vs source distribution

From this data set, we select a pair of receivers 158 m apart with a midpoint that coincides with the center of the ring in Figure 2.1. This ensures the same scattering regime for both receivers

at a given recording time. Integrating over a line surrounding the receiver pair is theoretically sufficient to get the Green's function because in our case the medium is lossless (see final section), but we select sources inside a ring to increase the signal-to-noise ratio (S/N) when computing the crosscorrelation. We define the azimuth with respect to the receiver pair, denoted by θ , as the difference between the azimuth defined by one source and the receiver pair center, on one hand, and the receiver pair azimuth on the other (Figure 2.2). The goal is to compare the convergence of the

Figure 2.2: The two geometries we considered, based on θ . The first is $|\theta| < 20^\circ$ and $180 - |\theta| < 20^\circ$, for which sources are inside the EFLs of the receiver pair. the second is $150^\circ > |\theta| > 30^\circ$, for which sources are outside the endfire lobes. In both cases, radii $R1$ and $R2$ are changed to 200–515 m and 300–400 m, respectively, to include the same number of sources as in Figure 2.1a.



two-point correlation function to the actual Green's function for different sets of source locations and different amounts of scattering governed by the selected time window. In fact, we consider two time windows. We study the contribution of direct arrivals to the correlation function in the 0–1.5-s interval and the contribution of coda waves in the 1.5–3-s time window.

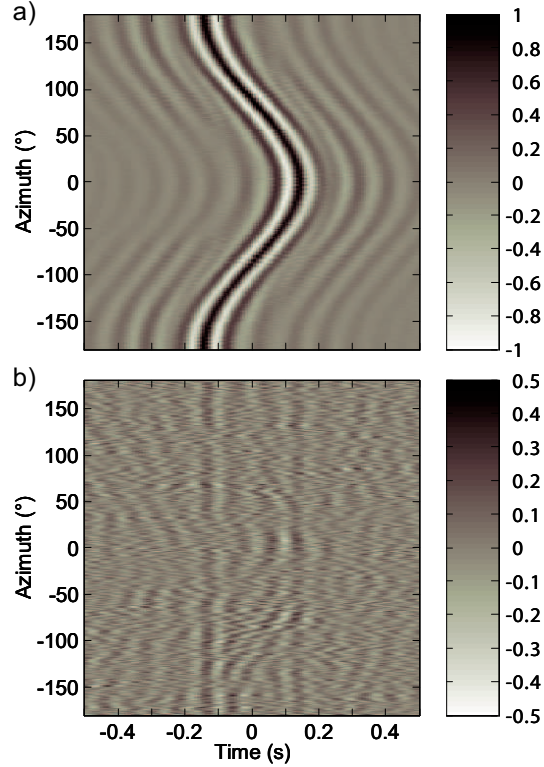
No processing is performed on the signals, such as frequency whitening or one-bit amplitude-only time-domain equalization. However, to cancel the intensity decay with time in the coda, the logarithmic decrease of the record amplitudes at times greater than 1.5 s is compensated.

We define the endfire lobes (EFLs) of the receiver pairs as the areas located in the alignment of the receivers (one on each side), with an aperture depending on the ratio between wavelength and range between receivers (Figure 2.2) [Roux & Kuperman, 2004]. We define the aperture of the EFLs as the misalignment of the source with respect to the receiver pair, that causes a delay between the apparent traveltimes between the two receivers and the actual Green's function traveltimes of $1/8$ of a period. In other words, the EFLs are areas in which the phase of the correlation function of direct waves is stationary with respect to azimuth θ . Because the angular aperture of the areas with sources inside or outside the EFLs is different (Figure 2.2), we adjust the $R1$ and $R2$ radii of the two circles to keep the same number of sources in each area. By doing so, we get comparable S/N in the correlation function, which is more convenient for comparing results.

To improve averaging over sources, and thus S/N of the two-point correlation function, and to ensure a laterally homogeneous medium, we perform an azimuthal average. This is done by considering 17 other receiver pairs, centered on the same point and spaced by the same distance but aligned in other directions θ . For each receiver pair, $R1$ and $R2$ remain the same but θ is redefined according to the receiver pair direction; the set of sources located inside the EFLs is updated.

In Figure 2.3, we compute the crosscorrelation function between signals from each source of Figure 2.1a and for two time windows (0–1.5 s and 1.5–3 s). We use a normalized version of the crosscorrelation such that the crosscorrelation $s_{SA} \times s_{SB}$ between two traces s_{SA} and s_{SB} recorded

Figure 2.3: Crosscorrelation function obtained from each source and plotted as a function of azimuth θ for two different time windows, **a.** 0–1.5 s and **b.** 1.5–3 s. Note the difference in the shading between the two figures.



for the same sources S at two receivers A and B is divided by the square root of the traces' energies E_{SA} and E_{SB} :

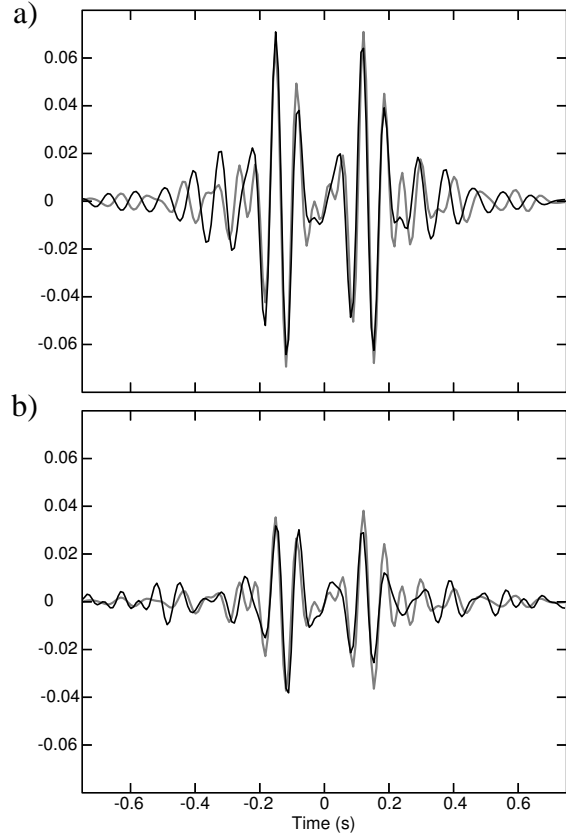
$$C_{AB}^S(t) = \frac{s_{SA} \times s_{SB}}{\sqrt{E_{SA} E_{SB}}}(t) \quad (2.1)$$

The amplitude of the crosscorrelation is thus a correlation coefficient as a function of the time shift. When using direct waves, the crosscorrelation function is highly dependent on source position, whereas the crosscorrelation is independent of source position when using scattered waves. However, the S/N of the crosscorrelation function is much lower for scattered waves than for direct waves. As a reference signal (i.e., the actual Green's function) for a given receiver pair (A,B) we used the closest source S_B to B recorded by the receiver in A ($G_{BA}(t)$) for positive times and the closest source S_A to A recorded in B ($G_{AB}(t)$) for negative times. To improve the comparison between this reference signal and the two-point correlation function $C_{AB}(t)$, we modified the time axis slightly to correct for the difference between the distances AB and AS_B as well as the difference between AB and S_AB . This 12.5-m distance correction is small compared to the 80-m central wavelength. The amplitude of the reference signal for negative times was also reversed because the relation we expect when averaging the crosscorrelation over sources is [Gouédard et al., 2008b]

$$-\frac{d}{dt} \sum_S C_{AB}^S(t) \propto G_{BA}(t) - G_{AB}(-t) \quad (2.2)$$

Figure 2.3a shows that when considering direct arrivals, the crosscorrelation function is highly dependent on the source azimuth. When stacking contributions of all sources (that is, all azimuths), all phases are averaged and only contributions of sources inside the EFLs do not vanish. The normalized crosscorrelation is computed independently for each source and the resulting functions are then averaged. It results in the function shown in Figure 2.4a. For the later time window, we are

Figure 2.4: Reconstructed Green's functions for the two time windows, **a.** 0–1.5 s and **b.** 1.5–3 s, when averaging correlation functions of Figure 2.3 (black lines). The gray lines are reference Green's functions, scaled to the amplitude of the crosscorrelation function for better comparison. Both direct and coda waves yield very good reconstructions of the Green's function.



in the scattering regime, and the crosscorrelation (Figure 2.3b) is noisy. Nevertheless, we see direct arrivals common to all crosscorrelation functions, around ± 0.1 s. When averaged over all sources, the S/N increases and direct-wave arrivals are quite evident in Figure 2.4b. Figure 2.4 shows that in both cases, the averaged crosscorrelation function is very close to the actual Green's function plotted in gray.

We crosscorrelate traces from two sets of sources (inside/outside the EFLs; see Figure 2.2) and using two time windows (0–1.5 s for direct waves and 1.5–3 s for coda waves; see Figure 2.1b). As we have seen when using direct waves, only contributions of sources located inside the EFLs do not vanish in the averaging process. Thus, considering only those sources should yield the same result, except that the S/N will be better because there is no competition between contributing and noncontributing sources. This is the result shown in Figure 2.5a. Regarding the opposite function, considering only sources outside the EFLs gives a correlation function that is not the Green's function (Figure 2.5c). Indeed, the peak that emerges in the correlation function is a consequence of the azimuthally dependent source selection because averaging does not produce a constructive interference when sources are outside the EFLs. This result clearly indicates the limitation of using the correlation of direct waves to extract the Green's function.

When using coda waves, the azimuthal distribution of incoming waves is no longer controlled by the distribution of sources but by the distribution of scatterers in the medium. When the distribution of scatterers is spatially uniform, coda waves naturally tend to become isotropic. This is illustrated by the correlation functions presented in Figure 2.5b and d. In Figure 2.5b, only coda waves produced by sources inside the EFLs are considered. A good agreement between correlation and actual Green's function is reached. The amplitude of the correlation function is lower than for

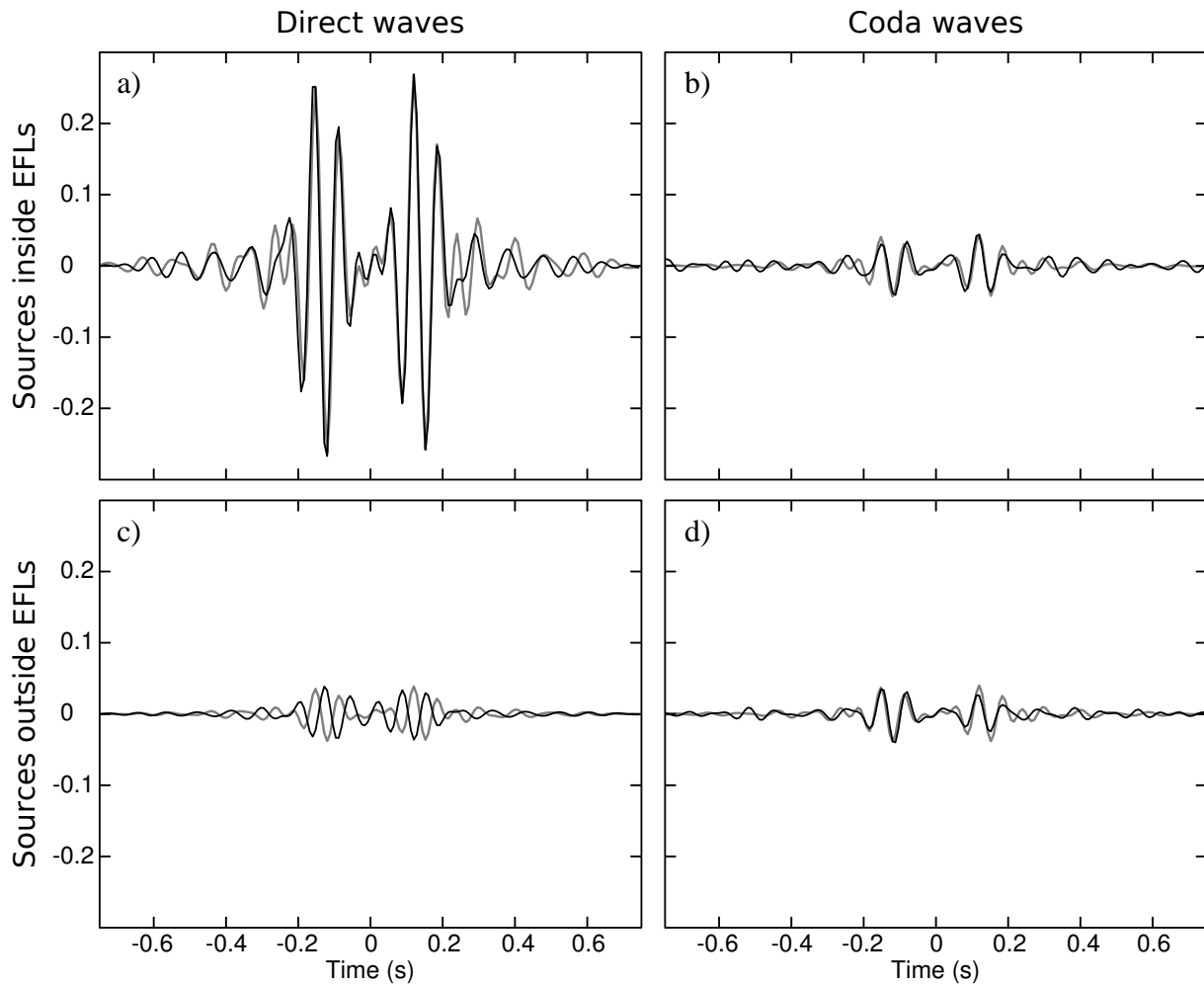


Figure 2.5: Crosscorrelation functions averaged over all sources (black line) and reference trace (gray line) for early and late time windows and different source distributions: **a.** sources inside the EFLs, 0–1.5 s; **b.** sources inside the EFLs, 1.5–3 s; **c.** sources outside the EFLs, 0–1.5 s; **d.** sources outside the EFLs, 1.5–3 s. In each plot, the reference trace is scaled to the amplitude of the crosscorrelation function for better comparison. When using direct waves (early time window), source locations are important and the Green's function is retrieved only if sources lay inside the EFLs. On the opposite, when using the scattered waves of the coda (late time window), source locations have no influence and the result is similar for both source distributions.

direct waves because only the scattered waves of the coda that travel between the two receivers contribute coherently to the Green's function.

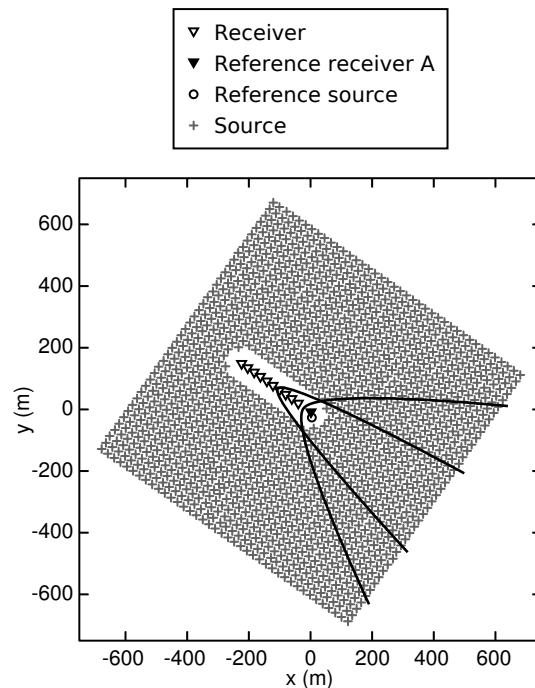
More spectacular is the case shown in Figure 2.5d, for which coda waves only from sources outside the EFLs are considered. The Green's function is nicely reconstructed, as for sources inside the EFLs. Note that the amplitude of the correlation function is independent of source positions inside/outside the EFLs when considering coda waves. This demonstrates the weak sensitivity of local coda waves to the source location, an expected feature for a wavefield evolving toward diffusion. These results confirm the role and importance of scattered waves in the convergence of the correlation function toward the Green's function. This also explains why practical applications are possible even with uneven spatial source distributions.

► 2.2.5 Amplitude of the crosscorrelation function

A question that arises about Green's function reconstruction using correlations is how relative amplitudes are retrieved when correlating records from receivers separated by different ranges. As seen, the answer to this question depends on the source distribution and the selected time window in the correlation process. The data set we analyze offers an opportunity to investigate this issue.

Figure 2.6 shows the experimental geometry used in this section. We select a line of 10 geo-

Figure 2.6: Experimental geometry considered for amplitude measurements. A line of 10 geophones is correlated with one geophone used as reference (A). The two grey hyperbola indicate the EFL-boundaries for two receiver pairs, the smallest hyperbola aperture corresponding to the largest offset between receivers and the largest aperture to the smallest offset.



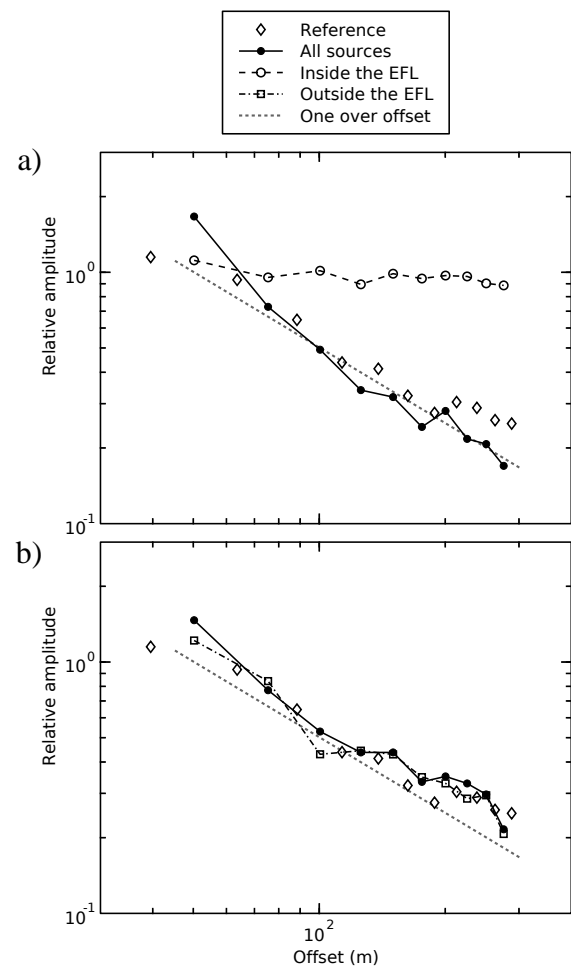
phones separated by 25 m and another geophone as reference (A) at 50 m, aligned with the geophone line. Correlation functions between the reference and each receiver of the linear array provide the reconstructed Green's function at 10 different ranges from 50 to 275 m. To improve the S/N of the correlation functions, we repeat the process 64 times with the network laterally shifted in both directions, keeping the same experimental interelement geometry. We then measure the amplitude (defined as the maximum of the envelope) of each of the 10 resulting correlation functions. The amplitude-versus-range measurements are compared to the amplitude of records made at each

receiver for the closest source to A. For this reference curve, the experimental decay is proportional to the inverse of the source-receiver range. This is from geometric spreading of the surface waves in the near-field regime (here, the wavelength is about 70 m and the maximum source-receiver distance is 275 m).

We measure the amplitude-versus-range curve of the correlation function for different sets of sources (only inside EFLs of each reference/receiver pair, only outside these EFLs, or for all available sources) and two different time windows (0–1.5 s for direct waves or 1.5–3 s for coda waves). Note that EFLs are defined with respect to the actual distance between receivers [Roux & Kuperman, 2004]; thus, we consider different sets of sources for each pair, as shown in Figure 2.6.

Figure 7 shows that relative amplitude with range strongly depends on source distribution and scattering. When considering all sources of the acquisition grid, the amplitude decay with range

Figure 2.7: Amplitude-versus-range measured on crosscorrelation functions from different source distributions (all sources as in Figure 2.6; sources inside the EFLs only or sources outside the EFLs only) and two time windows: 0–1.5 s corresponding to direct waves in **a** and 1.5–3 s corresponding to coda waves in **b**. Diamonds are amplitudes measured in the real data from the reference source. Dashed gray lines correspond to a one-over-offset decay.



is similar to the reference curve for both time windows. If we consider direct waves and sources located inside the EFLs only, the amplitude of the correlation does not depend on the range between receivers, as shown in Figure 2.7a. This can be explained using a time-reversal analogy [Derode et al., 2003b]. With the reciprocity principle, the sources inside the EFLs define a limited-aperture time-reversal mirror from which only wave vectors oriented in the direction of the receiver array are reconstructed. Thus, the phase of the correlation function is correct along the receiver array, but the amplitude corresponds to a plane wave. Using a distribution of sources inside EFLs only could lead to a measure of the medium attenuation because no geometric spreading is expected

on the amplitude-versus-range correlation function. Figure 2.7 shows that attenuation is negligible in the frequency bandwidth of interest in the recorded data. When using coda waves, the source distribution has no influence because waves are scattered. Thus, the correlation function provides the amplitude decay of the reference curve for sources located either inside or outside the EFLs.

► 2.2.6 Conclusion

The crosscorrelation of the seismic motions recorded at two points could yield the Green's function between these points. This result is only valid for particular source positions and/or wavefield properties. The data set we use allows us to investigate different cases by comparing the estimated Green's function obtained using correlation with the actual Green's function using different sets of source locations and different time windows. When using direct waves, source positions govern the convergence of the correlation function toward the Green's function. For coda waves, the source distribution has little influence, and the Green's function can be retrieved in any case. In the reconstructed Green's function, not only the phase but also relative amplitudes for receivers at different offsets are retrieved. This last result is true only for an isotropic wavefield, resulting from an even distribution of sources or from multiply scattered waves. When the source distribution is limited to EFLs, the correlation process only reconstructs a plane wave propagating in the receiver direction; amplitude decay with range no longer includes geometric spreading.

► 2.2.7 Acknowledgments

This work was part of a collaboration between Shell and University Joseph Fourier, Grenoble, initiated by the late Gérard Herman. His support and his inspiration are greatly acknowledged. The authors thank the Ministry of Oil and Gas of the Sultanate of Oman and Shell Research for permission to publish these results.

2.3 Imagerie passive d'un milieu homogène

► 2.3.1 Analyse fréquence/nombre d'onde (FK)

Les méthodes d'imagerie du sous sol par les ondes de surface reposent sur la mesure de courbes de dispersion. Ces courbes de dispersion doivent alors être identifiées, c'est-à-dire attribuées à tel ou tel mode des ondes de Rayleigh ou de Love. Cette identification n'est pas toujours triviale, en particulier lorsque des modes ont des vitesses proches ou lorsqu'ils ne sont pas tous excités (on peut par exemple rencontrer des cas où le mode fondamental n'est pas visible, mais le premier mode harmonique l'est). Une fois les courbes identifiées, elle sont « inversées » afin d'obtenir le profil de vitesse en fonction de la profondeur. Les algorithmes d'inversion sont souvent basés sur le calcul du problème direct dans un modèle initial que l'on fait évoluer de façon itérative afin de le faire correspondre aux courbes de dispersion mesurées. Il existe aussi des inversion de type *Monte Carlo* basées sur la convergence statistique, mais elle sont souvent coûteuses en temps de calcul.

Il existe différentes méthodes de mesure de courbes de dispersion. La plus simple consiste à filtrer les enregistrements dans une bande étroite autour d'une fréquence donnée, et à mesurer le temps de parcours de l'onde entre la source et le récepteur pour cette fréquence. Connaissant la distance source/capteur, on en déduit la vitesse de propagation des ondes. On recommence alors la mesure pour toutes les fréquences. Cette méthode a l'avantage d'être très simple et intuitive, mais elle ne permet pas de prendre en compte la cohérence spatiale de l'onde au cours de la propagation. En effet, si on mesure le champ d'onde issu d'une source en différents points de l'espace, on peut prendre en compte la mesure effectuée à une distance de la source pour introduire un *a priori* sur la mesure à d'autres distances. C'est d'une certaine façon ce que proposent les méthodes basées sur l'analyse fréquence/nombre d'onde, en projetant l'information mesurée dans un espace temps/distance dans les espaces réciproques pour chacune des deux variables.

Dans ce manuscrit, les techniques utilisant l'espace fréquence/nombre d'onde (ou FK) seront abordées trois fois : dans ce paragraphe et dans les paragraphes 3.1.2 et 5.1.1. J'ai cependant voulu séparer les trois applications car, même si l'algorithme utilisé est très proche, voir rigoureusement identique, les approches et les justifications des méthodes sont très différentes. Dans ce paragraphe, l'objectif est de mesurer une courbe de dispersion à partir d'une section sismique obtenue grâce à un profil linéaire, en utilisant un algorithme linéaire très simple. Dans le paragraphe 3.1.2, l'enjeu est aussi de mesurer une courbe de dispersion, mais à partir de bruit de fond enregistré sur des réseaux 2D, et en utilisant des méthodes plus complexes pouvant être non linéaires. Le paragraphe 5.1.1, quant à lui, traitera de la localisation de sources dans un milieu de propagation supposé connu.

Les méthodes FK sont basées sur une décomposition du champ en ondes planes. Une telle onde se propageant le long d'une droite d'abscisse curviligne r peut être décrite (à un terme d'amplitude près) dans le domaine de Fourier par :

$$s(\omega, r) = \exp\left(i \frac{\omega r}{c_0(\omega)}\right) = \exp(ik_0(\omega) r) \quad , \quad (2.3)$$

où $c_0(\omega)$ est la vitesse de phase de cette onde pour la pulsation ω , et $k_0(\omega)$ le nombre d'onde associé. La transformée de Fourier spatiale de ce signal s'écrit

$$\hat{s}(\omega, k) = \int \exp(ik_0(\omega) r) \exp(-ikr) dr \quad (2.4)$$

$$= \int \exp(i(k_0 - k)r) dr \quad . \quad (2.5)$$

$\hat{s}(\omega, k)$ est nulle pour tout k différent de k_0 . Son calcul pour tout k (à ω fixé) permet ainsi de retrouver le vecteur d'onde $k_0(\omega)$, et donc la vitesse de phase $c_0(\omega)$. On fait alors varier ω pour retrouver la courbe de dispersion complète.

Dans la pratique, lorsque plusieurs ondes sont présentes dans le champ d'onde enregistré, on trouve plusieurs vecteurs d'onde donnant une valeur de \hat{s} non nulle pour une même pulsation, comme c'est le cas sur les figures 2.9b et 2.9d ou encore sur la figure 4.4 un peu plus loin. Il convient alors de les interpréter proprement (mélange d'ondes P et de Rayleigh, différents modes d'une même onde, ...) afin de pouvoir utiliser ces mesures dans un algorithme d'inversion.

Le cas des échantillonnages spatiaux irréguliers peut poser problème pour le calcul de la transformée de Fourier. En effet, l'algorithme de transformée de Fourier rapide (*fft*) largement utilisé

pour les calculs numériques ne permet pas de traiter ces cas. Ce problème est cependant souvent rencontré en géophysique où la configuration du terrain ne permet pas toujours de garder un espacement régulier entre les récepteurs. Les théories traitant de la transformée de Fourier de signaux irrégulièrement échantillonnés sont variées, chacune ayant ses spécificités. La pertinence de chacune de ces théories est généralement fonction de l'échantillonnage effectif. On comprend bien que lorsque seuls quelques points de mesure manquent pour obtenir un échantillonnage régulier, et que le pas est petit devant la longueur d'onde recherchée (cas d'un récepteur défectueux par exemple) il est possible de retrouver la trace manquante à partir d'une interpolation basée sur les traces voisines. Ce type d'approche devient dangereux pour des échantillonnages très irréguliers où le nombre de traces à interpoler afin d'obtenir un échantillonnage régulier est plus important que le nombre de points de mesure effectifs.

La méthode la plus simple pour traiter le problème, et qui est celle que j'ai employée ici, consiste à calculer la transformée de Fourier discrète (ou *dft*) selon la formule suivante :

$$\text{dft}[f(x)](k) = \sum_{n=1}^N f(x_n) e^{ikx_n} , \quad (2.6)$$

où x_n , $1 \geq n \geq N$ sont les points d'échantillonnage spatial et k est le vecteur d'onde. Cette équation est une transcription numérique naïve de l'équation algébrique définissant la transformée de Fourier. Elle a l'avantage de permettre de choisir les nombres d'onde k pour lesquels la transformée de Fourier sera calculée, ce qui n'est pas le cas avec l'algorithme de transformée de Fourier rapide. Cette méthode est efficace tant que la densité spatiale de points de mesure est à peu près homogène, c'est-à-dire qu'il n'y a pas de zones sous-échantillonnées ou à l'inverse sur-échantillonnées.

Lors d'utilisations de données réelles, le nombre de points d'échantillonnage en espace est souvent petit devant le nombre de points en temps. La résolution dans le domaine des nombres d'onde est alors faible devant la résolution en fréquences. Le moyen habituellement utilisé pour augmenter la résolution en nombre d'onde est le remplissage par des zéros (ou *zero padding* en anglais). La distance maximale enregistrée est ainsi artificiellement augmentée en ajoutant des traces nulles correspondant à de grandes distances. Cette méthode, très souvent utilisée puisque très facile à mettre en œuvre, présente l'inconvénient de rajouter une information *a priori* fausse puisque non physique. J'ai utilisé une méthode qui permet de doubler le nombre de points d'échantillonnage spatial en étendant la section sismique vers les temps et les distances négatifs par symétrie. Ainsi, chaque trace considérée dans les temps positifs pour les distances positives sera « recopiée » dans les temps négatifs pour les distances négatives. Cette information reste juste puisque le milieu étant réciproque la propagation est identique dans les deux sens. Lorsque cela ne suffit pas à avoir une résolution suffisante en nombre d'onde, le remplissage par des zéros est ensuite utilisé, vers les grandes distances positives et négatives.

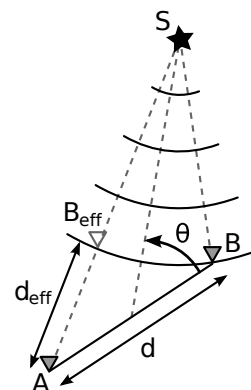
Cette méthode permet d'améliorer le résultat tant que l'échantillonnage spatial est satisfaisant pour les distances faibles. Lorsque ce n'est pas le cas, elle peut introduire des biais dans l'estimation du spectre. En effet, si on considère un échantillonnage régulier dans une plage de distances d_{min} – d_{max} , avec d_{min} plus grand que le pas d'échantillonnage spatial, alors symétriser le signal vers les temps et distances négatifs introduit un défaut d'échantillonnage autour de zéro, ce qui a pour effet de biaiser la transformée de Fourier en la convoluant par la transformée de Fourier d'une double porte.

► 2.3.2 Application de l'analyse FK aux données Qarn Alam

J'ai utilisé la méthode décrite dans le paragraphe précédent afin de mesurer les courbes de dispersion des ondes de surface pour les données Qarn Alam. Afin d'utiliser au mieux la multitude de sources enregistrées, j'ai exploité l'hypothèse du milieu homogène afin de réaliser une projection permettant d'exploiter les fonctions de corrélation entre toutes les paires de récepteurs, pour toutes les sources.

Nous avons vu dans le paragraphe 2.2.4 que lorsqu'on construit la fonction de corrélation moyennée sur les sources entre deux points A et B, seules certaines sources contribuent à la reconstruction de la fonction de Green, tandis que d'autres s'annulent de proche en proche dans la moyenne sur les sources. Les sources qui contribuent de façon cohérente à la reconstruction sont celles placées dans l'alignement des deux points, avec un écart à l'alignement autorisé dépendant du rapport entre la distance inter capteurs et la longueur d'onde (ce qui définit les *endfire lobes*). Si maintenant on considère le cas d'une source S située hors de cet alignement, dans le cas d'un milieu homogène, les choses peuvent être interprétées différemment. Dans un tel milieu, les fronts d'ondes sont cylindriques, et ce qu'on enregistre à un récepteur B est ce qu'on aurait enregistré tout au long du front d'onde passant par B (Figure 2.8). On peut donc définir une position effective du récepteur B, notée B_{eff} , pour laquelle la source est dans l'alignement de la paire AB_{eff} . B_{eff} est l'image du point B par la projection circulaire de centre la source S sur la droite définie par A et S. En effectuant la corrélation entre les enregistrements en A et B pour cette source, on obtient la fonction de Green pour la distance effective $d_{\text{eff}} = d(AB_{\text{eff}})$ (dans un milieu homogène, la fonction de Green ne dépend que de la distance entre les récepteurs et non de leur position). La distance d_{eff} pour une paire de récepteurs est définie par rapport à l'azimut θ de la source. Ainsi, en considérant seulement deux récepteurs et une collection de sources les entourant (*i.e.*, incluant tous les azimuts), on peut construire une section sismique avec des distances d_{eff} réparties entre 0 (pour $\theta = \pm 90^\circ$) et la distance réelle $d = d(AB)$ entre les récepteurs (pour $\theta = 0$ ou $\pm 180^\circ$).

FIG. 2.8 – On définit l'azimut θ de la source S par rapport à la paire de récepteurs AB. On définit également une position effective B_{eff} du récepteur B telle que la source S soit alignée avec AB_{eff} . L'intercorrélation entre les enregistrements en A et en B pour cette source donne la fonction de Green pour la distance effective $d_{\text{eff}} = d(AB_{\text{eff}})$.



Nous avons appliqué ce principe au jeu de données d'Oman. L'approximation de milieu homogène n'est pas valable dans ce cas, mais elle nous permet d'obtenir un modèle de vitesse homogène moyen pour le milieu. En considérant les 1600×1599 paires de récepteurs, et les 1600 sources, nous avons calculé les $4 \cdot 10^9$ fonctions de corrélation possibles, ainsi que les distances effectives d_{eff} correspondantes. Les fonctions de corrélation ayant la même distance effective (à un mètre près, pour une longueur d'onde de l'ordre de 100 m) sont ensuite moyennées, quelque soit la paire de capteurs correspondante (le milieu étant supposé homogène, la position des récepteurs n'a pas

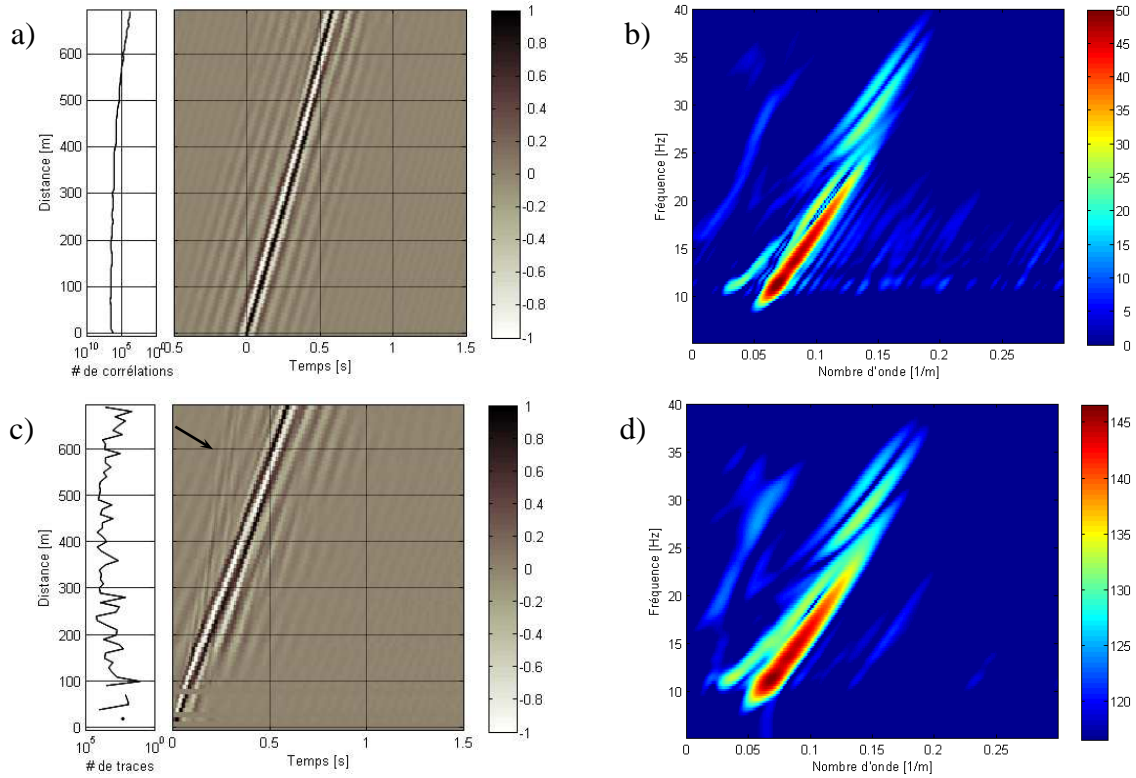
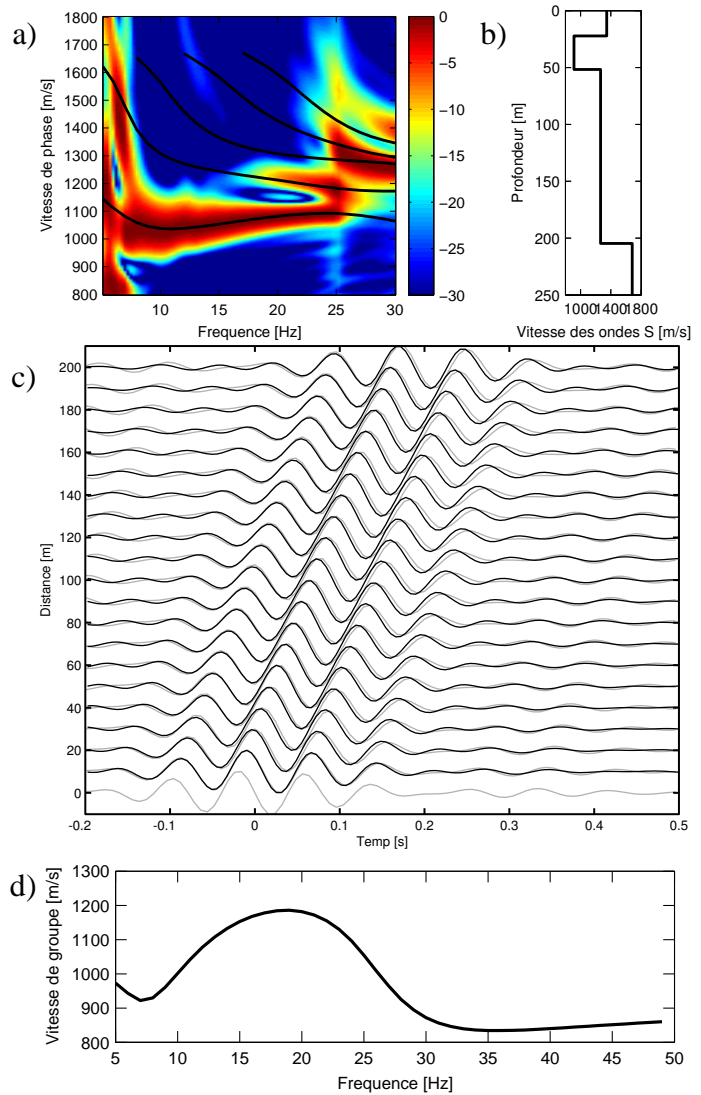


FIG. 2.9 – **a.** La section sismique obtenue à partir des corrélations dans l'approximation de milieu homogène. Les ondes de Rayleigh sont majoritairement reconstruites. Le nombre de fonctions de corrélation moyennées pour chaque distance est indiqué sur la gauche de la figure. Les traces à chaque distance sont normalisées sur la figure pour faciliter la comparaison. **c.** Le même type de moyenne effectué sur les données brutes, avec le nombre de traces moyennées précisé sur la gauche. **b.** et **d.** Les diagrammes FK (en échelle dB) correspondant aux deux sections sismiques. Les deux diagrammes sont très semblables, mais avec une différence dans la dynamique d'amplitudes.

d'influence). On obtient ainsi la section sismique de la figure 2.9a. Dans cette section, les ondes de Rayleigh sont majoritaires, comme attendu au vu de la géométrie du réseau. A titre de comparaison, nous avons réalisé le même type de moyenne à partir des données brutes : pour chaque récepteur, nous avons moyenné tous les enregistrements issus de sources situées à la même distance (ici encore, à un mètre près), puis nous avons moyenné les sections sismiques ainsi obtenues pour chaque récepteur. La section sismique ainsi obtenue est celle de la figure 2.9c. Il est important de noter la différence d'ordre de grandeur dans le nombre de traces moyennées pour chaque distance pour les figures 2.9a et 2.9c. En effet, la section sismique de la figure 2.9c met en jeu 1600 fois moins de traces que celle de la figure 2.9a.

Sur la section sismique obtenue à partir des données brutes, on voit clairement une onde rapide (indiquée par la flèche sur la figure 2.9c). Il semble que ce soit une onde réfractée, mais je n'ai pas la connaissance du milieu nécessaire à son identification précise. Cette onde n'est pas visible sur la figure 2.9a. Pour une comparaison plus précise entre ces deux sections sismiques, j'ai appliqué à ces deux sections un traitement fréquence-nombre d'onde (FK), dont les résultats sont présentés dans les figures 2.9b et 2.9d. Les deux diagrammes FK sont très semblables, sauf en ce qui concerne leur dynamique en amplitude : la figure 2.9b a une dynamique de 50 dB, tandis que celle de la figure 2.9d n'est que de 30 dB. La résolution est meilleure dans le diagramme obtenu par corrélation, grâce

FIG. 2.10 – **a.** Courbes de dispersion mesurées à l'aide de la section de corrélation. La figure est normalisée pour chaque fréquence, et l'amplitude est en dB. Les courbes noires présentent les 5 premiers modes de Rayleigh calculés à partir du modèle obtenu par inversion. **b.** Modèle de vitesse des ondes S obtenu par inversion des courbes de dispersion. **c.** Traces synthétiques (en noir) obtenues pour le modèle de vitesse, à comparer à la section de corrélation (en gris). Les sismogrammes à chaque distance sont normalisés sur la figure. **d.** Courbe de dispersion en vitesse de groupe du mode fondamental dans le modèle présenté en b.



à un meilleur échantillonnage spatial, en particulier pour les faibles distances, ainsi qu'à un très bon rapport signal sur bruit. L'onde rapide de la figure 2.9c correspond à la région autour du point ($0,05 \text{ m}^{-1}, 30 \text{ Hz}$) de la figure 2.9d. Une zone similaire est visible sur la figure 2.9b, ce qui indique que cette onde rapide est aussi présente dans la section sismique de la figure 2.9a, mais avec une amplitude beaucoup plus faible que l'onde de Rayleigh, si on considère la différence de dynamique entre les deux diagrammes. On peut, à partir du résultat de la figure 2.9b, sélectionner la région correspondant à cette onde dans le plan FK, et réaliser la transformée inverse pour revenir dans le plan temps/distance. On isole ainsi cette onde rapide dans la section de corrélations, et j'ai vérifié qu'elle se superpose parfaitement à celle de la section brute. La section sismique reconstruite par corrélation inclue toute la complexité du milieu, et elle peut donc être utilisée dans des processus d'inversion à des fins d'imagerie pour retrouver un modèle moyen homogène. J'ai donc mesuré les courbes de dispersion des ondes de Rayleigh à partir de la section corrélée, et j'ai ensuite inversé ces courbes afin d'obtenir le modèle de vitesse présenté sur la figure 2.10b.

Le modèle obtenu est un modèle homogène équivalent. Il explique bien les courbes de dispersion, au moins pour les deux premiers modes de Rayleigh, comme le montre la figure 2.10a. À partir du second mode harmonique, les modes sont proches et les courbes de dispersion ne peuvent être mesurées dans les données à cause des interactions entre modes. J'ai généré des sismogrammes

synthétiques à partir du modèle obtenu par inversion, pour les mêmes distances que celles de la section de corrélation, à partir d'un code de calcul reproduisant la solution exacte de l'équation de propagation (AXITRA, O. Coutant, 1989). La figure 2.10c montre que ces traces synthétiques (en noir) reproduisent bien la section de corrélation (en gris).

La figure 2.9a a été réalisée en utilisant les 1600 sources disponibles pour la moyenne, mais le nombre de sources peut être diminué étant donné que la corrélation multiplie le nombre de traces disponibles pour la moyenne. J'ai ainsi construit des sections sismiques similaires en utilisant 100 sources (1 sur 4 dans les deux directions du plan), puis seulement 4 (une à chaque coin du réseau), et enfin une seule source située dans un coin du réseau. La totalité des 1600 récepteurs disponibles est utilisée dans chaque cas. Les sections reconstruites sont présentées sur la figure 2.11, ainsi que les diagrammes FK correspondants. On voit sur cette figure que la réduction du nombre de sources utilisées de 1600 à 100 ne change pas le résultat et que la résolution dans le diagramme FK est encore très bonne. Lorsque l'on n'utilise que 4 sources pour former la section de corrélation, le diagramme FK commence à se dégrader, et en particulier l'onde rapide de la figure 2.9 n'est plus aussi clairement identifiable. Les ondes de Rayleigh sont toutefois encore visibles, et une courbe de dispersion pourrait être mesurée à partir de ce résultat. Enfin, lorsqu'une seule source est utilisée, la section sismique et le diagramme FK sont bruités. Ils restent cependant exploitables même si la précision diminue. Ces résultats sont une illustration du fait que dans les techniques de corrélation la résolution est gouvernée par le réseau de capteurs et non celui de sources. En effet, ce qui permet d'avoir une mesure encore exploitable à partir d'une seule source est le nombre important de récepteurs, et donc de couples pour la corrélation ($1600 \times 1599 \approx 2.56 \cdot 10^6$ dans notre cas). Le grand nombre de couples permet alors de réduire considérablement le bruit dans la section sismique grâce à la moyenne sur les corrélations, et ainsi de faire ressortir l'information utile, même lorsque celle-ci est de faible amplitude.

2.4 Tomographie passive d'ondes de surface

Afin d'obtenir une image plus réaliste du milieu, les hétérogénéités latérales du milieu doivent être prises en compte, et l'hypothèse de milieu homogène abandonnée. La tomographie d'ondes de surface a pour but d'imager ces hétérogénéités de proche surface. Elle nécessite habituellement de nombreuses sources sismiques pour échantillonner le milieu de façon satisfaisante. D'un autre côté, les méthodes de corrélation permettent de « créer » des données en reconstruisant la fonction de Green entre les récepteurs d'un réseau à partir d'enregistrements de champ diffus. Nous avons voulu montrer ici la faisabilité de la tomographie d'ondes de surface à partir des fonctions de Green reconstruites par corrélation. Le travail présenté ici ne va pas jusqu'à l'obtention du modèle de vitesse mais s'arrête à la mesure des courbes de dispersion en fonction de la position dans le réseau. Je parlerai donc d'inversion et non de tomographie. J'ai utilisé un algorithme d'inversion des temps de parcours des ondes de Rayleigh afin d'obtenir un modèle de vitesse de ces ondes pour le milieu. La méthode utilisée est une méthode de quasi-Newton basée sur les travaux de Tarantola & Valette [1982] et les développements de Tarantola [1987]. Le problème considéré ici étant un cas simple de cette méthode, l'inversion est mise en œuvre en une seule itération. Sans entrer dans les

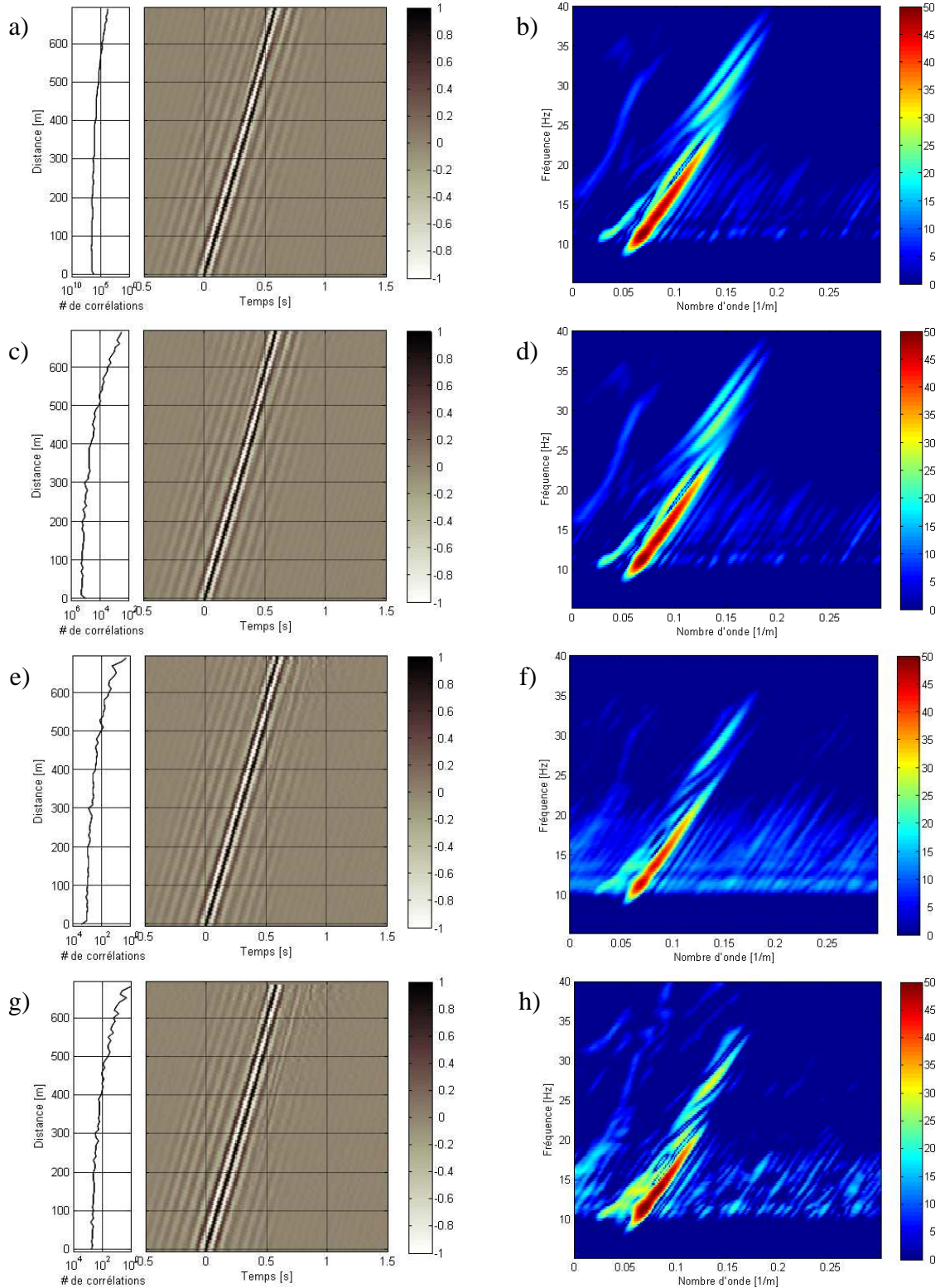


FIG. 2.11 – Sections de corrélation et diagrammes FK obtenus en utilisant de moins en moins de sources pour construire la moyenne (mais toujours les 1600 récepteurs disponibles). La première ligne (a et b) est obtenue pour 1600 sources (figures identiques à celles présentées en 2.9a et b) ; la deuxième ligne pour 100 sources (une sur 4 dans les deux directions du plan) ; la troisième pour 4 sources (une à chaque coin du réseau) ; la dernière ligne est obtenue en n'utilisant qu'une seule source située dans un coin du réseau. Pour chaque section, le nombre de fonctions de corrélation effectivement moyennées pour chaque distance est précisé sur la gauche.

détails de la théorie de l'inversion, je vais expliquer ici l'algorithme simplifié utilisé et les résultats obtenus à partir des données.

► 2.4.1 Algorithme d'inversion

On considère un jeu de données t_m constitué de N_t temps de parcours mesurés, correspondant à autant de trajets entre des éléments d'un réseau. On discrétise l'espace en N_c cellules. La méthode est dite Bayésienne, c'est-à-dire qu'elle suppose la connaissance d'un modèle de vitesse *a priori* pour la discrétisation de l'espace considérée, que l'on prendra comme modèle initial v_i pour l'inversion. Ce modèle initial, constitué de N_c éléments décrivant la vitesse des ondes dans chacune des cellules, permet de calculer un jeu de données initiales t_i pour les mêmes trajets que ceux de t_m . La méthode consiste à trouver la meilleure façon de faire évoluer le modèle v_i grâce à une perturbation Δv , de telle façon que les temps de parcours dans le modèle $v = v_i + \Delta v$, notés $t(v)$, soient le plus proche possible de t_m . On suppose que le modèle *a priori* est assez proche du modèle réel, de telle façon que Δv reste une perturbation faible. Afin de trouver Δv , on définit une fonction coût $S(v)$ qui traduit à la fois l'écart entre t_m et $t(v)$, noté $\Delta t = t(v) - t_m$, et l'écart au modèle *a priori* Δv . Afin de pondérer l'importance de ces deux grandeurs dans la fonction coût, on définit les matrices de covariance des données C_d et du modèle C_m , qui décrivent l'incertitude sur les données mesurées et sur le modèle *a priori*. La fonction coût est alors définie comme

$$S(v) = \frac{1}{2} \left(\Delta t C_d^{-1} {}^t\Delta t + \Delta v C_m^{-1} {}^t\Delta v \right) , \quad (2.7)$$

où ${}^t \cdot$ représente le transposé (Δt et Δv sont des vecteurs ligne). On cherche alors la perturbation Δv qui minimise la fonction coût, que l'on notera Δv_f . D'après l'équation (2.7), si $\|C_d\| \ll \|C_m\|$ (i.e., l'erreur sur les données est faible et le modèle *a priori* mal connu), la fonction coût est totalement contrôlée par Δt , et l'inversion tendra à trouver un modèle de vitesse expliquant parfaitement les données, mais qui pourra être très éloigné du modèle de vitesse *a priori*. Inversement, si $\|C_d\| \gg \|C_m\|$ (i.e., l'erreur sur les données est importante et le modèle *a priori* bien contraint), les données mesurées auront peu d'influence sur la fonction coût, et le modèle final sera proche du modèle *a priori* (i.e., Δv sera faible). On voit donc l'importance du rapport $\|C_d\|/\|C_m\|$, et j'ai défini Δv_{max} , écart maximal au modèle initial autorisé, et normalisé C_m de façon à le prendre en compte (voir plus bas l'équation (2.14) pour l'expression de C_m).

Afin de trouver le minimum de la fonction $S(v)$, on calcule la dérivée de l'équation (2.7) par rapport à v . Pour cela, il faut définir la matrice des dérivées partielles (ou dérivées de Fréchet) $G = \partial t / \partial v(v)$, de taille $N_t \times N_c$. G dépend du modèle v considéré. On a alors

$$\gamma(v) = \frac{\partial S}{\partial v}(v) = {}^tG(v) C_d^{-1} \Delta t + C_m^{-1} \Delta v . \quad (2.8)$$

Dans notre cas, la matrice G est facile à calculer puisque l'on connaît le modèle *a priori* et la position des éléments du réseau. En effet, si on note D la matrice de taille $N_t \times N_c$ qui représente, pour chaque trajet considéré, la distance parcourue dans chaque cellule, on a $G(v) = -D/v^2$. Afin de trouver le minimum de S , on cherche le modèle v qui annule γ . Comme on ne connaît G (et donc γ) que pour le modèle *a priori* v_i , on utilise un développement de γ autour de v_i :

$$\gamma(v) \approx \gamma(v_i) + \frac{\partial \gamma}{\partial v}(v_i) \Delta v . \quad (2.9)$$

$\partial\gamma/\partial v = \partial^2 S/\partial v^2$ s'appelle de Hessian de la fonction coût. Il peut être approximé en dérivant l'équation (2.8) et en négligeant les termes de second ordre (on parle alors de quasi-Hessian, d'où le nom de la méthode, dite de quasi-Newton) :

$$\frac{\partial\gamma}{\partial v}(v_i) \approx {}^tG(v_i) C_d^{-1} G(v_i) + C_m^{-1} \quad . \quad (2.10)$$

L'annulation de l'équation (2.9) donne alors la perturbation finale Δv_f permettant de minimiser la fonction coût :

$$\Delta v_f = - \left({}^tG(v_i) C_d^{-1} G(v_i) + C_m^{-1} \right)^{-1} \gamma(v_i) \quad (2.11)$$

$$= - \left({}^tG(v_i) C_d^{-1} G(v_i) + C_m^{-1} \right)^{-1} {}^tG(v_i) C_d^{-1} \Delta t_i \quad . \quad (2.12)$$

La matrice $\hat{G}^{-1} = - \left({}^tG(v_i) C_d^{-1} G(v_i) + C_m^{-1} \right)^{-1} {}^tG(v_i) C_d^{-1}$ est le pseudo-inverse de G . Elle peut s'exprimer de façon équivalente comme $\hat{G}^{-1} = -C_m {}^tG(v_i) \left(G(v_i) C_m {}^tG(v_i) + C_d \right)^{-1}$, la première formulation étant mieux adaptée aux cas sous-déterminés ($N_t > N_c$) et la deuxième aux cas sur-déterminés ($N_t < N_c$) pour des raisons de temps de calcul liées aux tailles des matrices C_d et C_m .

L'efficacité de cette méthode d'inversion repose sur une définition appropriée des matrices C_d et C_m . La matrice C_d est homogène au carré d'un temps. Elle permet de pondérer l'importance de chacune des mesures en précisant l'erreur faite sur chacune d'elle. Lorsque cette matrice n'est pas diagonale, elle rend compte de la dépendance des mesures entre elles. J'ai choisi de considérer une matrice diagonale dont tous les éléments sont égaux à la variance de $\Delta t_i = t(v_i) - t_m$:

$$C_d = I_{N_t} \frac{1}{N_t} \sum \Delta t_i^2 \quad , \quad (2.13)$$

où I_{N_t} est la matrice identité de taille $N_t \times N_t$.

La matrice C_m est homogène à une vitesse au carré. Elle permet de pondérer l'erreur commise pour chaque cellule dans la définition du modèle *a priori* v_i (ou d'un point de vu opposé, la confiance dans ce modèle). Elle est aussi utile pour lier les valeurs des vitesses entre les cellules. J'ai utilisé une matrice C_m ayant la même valeur sur la diagonale, c'est-à-dire que la même erreur est faite pour chaque cellule lors de la définition de v_i . Pour les éléments de P hors de la diagonale, j'ai choisi de lier les cellules en fonction de leur éloignement, ce qui introduit un lissage du modèle final. De plus, nous avons vu que le rapport $\|C_d\|/\|C_m\|$ permet de contrôler l'inversion, et j'ai donc introduit dans l'expression de C_m la variable Δv_{max} évoquée plus haut. La définition de C_m retenue est donc, pour l'élément (i, j) :

$$C_m(i, j) = \frac{\Delta v_{max}^2}{3} e^{-d(i, j)/d_0} \quad , \quad (2.14)$$

où $d(i, j)$ représente la distance entre les centres des cellules i et j , et d_0 est la longueur de corrélation du lissage introduit. Le facteur $\Delta v_{max}^2/3$ vient du fait que la covariance d'une variable aléatoire X de densité de probabilité uniforme entre $-\Delta X$ et ΔX est $\Delta X^2/3$.

Le modèle de vitesse final v_f est calculé à partir de la perturbation Δv_f obtenue à la fin de l'inversion selon l'équation (2.12). La diminution relative de la valeur de la fonction coût permet

d'évaluer l'efficacité de l'inversion. La qualité de l'inversion peut aussi être mesurée en estimant les temps de parcours $t(v_f)$ dans le modèle final, et en les comparant aux temps mesurés t_m . L'erreur résiduelle $\Delta t_f = t(v_f) - t_m$ résulte d'erreurs d'inversion, dues à une mauvaise discrétisation de l'espace par exemple, mais aussi aux erreurs de mesure des temps de propagation t_m . On peut minimiser l'influence des erreurs de mesure en redéfinissant une nouvelle mesure $t_m^1 = t_m + \text{perm}(\Delta t_f)$, où $\text{perm}(\Delta t_f)$ est un vecteur dont les éléments sont ceux de Δt_f , permutés de façon aléatoire. Cette opération permet de redistribuer les éventuelles erreurs de mesure entre les différents trajets de t_m . Le vecteur t_m^1 est alors utilisé pour calculer $\Delta t_i^1 = t_i - t_m^1$, que l'on utilise pour déterminer un nouveau modèle de vitesse v_f^1 . Cette opération, appelée *bootstrapping*, peut être répétée N_{bs} fois, en définissant de nouvelles permutations de Δt_f donnant de nouveaux vecteurs $t_m^1 \dots t_m^{N_{bs}}$. On obtient ainsi un ensemble de N_{bs} vecteurs $v_f^1 \dots v_f^{N_{bs}}$, qui définissent un ensemble de modèles acceptables expliquant les données t_m , aux erreurs de mesure près. Le résultat de l'inversion est alors le modèle de vitesse moyen (*i.e.*, la vitesse dans chaque cellule est égale à la moyenne des vitesses obtenues pour cette cellule pour chacune des itérations), auquel on peut associer une estimation de la précision égale à la variance calculée sur les N_{bs} modèles.

► 2.4.2 Inversion de la zone de Qarn Alam

L'algorithme d'inversion est appliqué aux données brutes dans un premier temps, afin de vérifier sa robustesse. Dans un second temps, les même traitement sera appliqué à des fonctions de Green reconstruites pour imager le même milieu. Les deux images seront comparées pour étudier la faisabilité de l'imagerie passive à ces échelles.

Utilisation données brutes

Nous avons travaillé dans la bande de fréquences 10-25 Hz. Les différentes étapes de l'inversion sont les suivantes :

1. on sélectionne un sous réseau définissant une cellule élémentaire, située par exemple dans le coin inférieur gauche de la zone à étudier (voir figure 2.12a). Cette cellule est un carré constitué d'une ligne de sources sur la gauche et d'une ligne de récepteurs sur la droite. Chaque ligne est composée de 8 éléments, ce qui donne un carré de 200 m de côté.
2. pour augmenter le rapport signal sur bruit des enregistrements, nous avons utilisé la ligne de sources pour former une onde plane se propageant de la gauche vers la droite dans la cellule élémentaire. Cette onde plane est enregistrée en chaque point récepteur, et les temps d'arrivée des ondes de Rayleigh sont mesurés (définis ici comme le temps du maximum de l'enveloppe). L'utilisation d'ondes planes permet également d'améliorer la mesure dans le cas des milieux inhomogènes, en favorisant les ondes directes par rapport aux ondes réfléchies. La mesure du temps de propagation sur l'enveloppe du signal impose un modèle de vitesse en vitesse de groupe. Afin d'obtenir un modèle de vitesse de phase il faudrait mesurer la phase de l'onde plane à au niveau de la ligne de récepteurs, ce qui ajoute une difficulté supplémentaire, cette phase n'étant connue qu'à 2π près.
3. en échangeant les positions des lignes de sources et de récepteurs, on construit de la même façon une onde plane se propageant de la droite vers la gauche, dont on mesure les temps

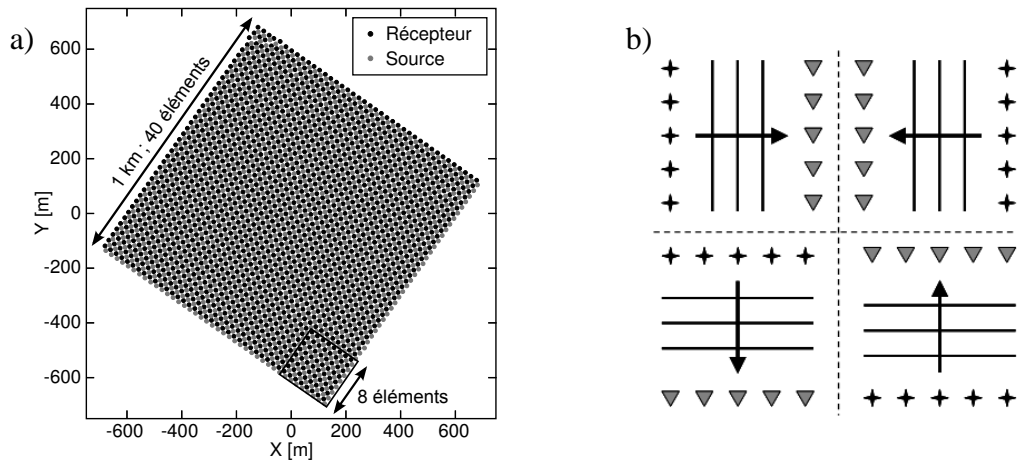


FIG. 2.12 – **a.** Le réseau utilisé couvrant la zone d'1 km carré. Le carré noir en bas à gauche donne la position de la première cellule élémentaire utilisée. **b.** Les cellule élémentaires sont constituées de lignes de sources (étoiles noires) et de récepteurs (triangles gris) disposées de 4 façons différentes afin de générer 4 ondes planes se propageant dans 4 directions (troisième étape du traitement).

de parcours jusqu'à chacun des récepteurs. Les deux ondes planes ainsi considérées échantillonnent le même milieu, mais constituent des mesures indépendantes des temps de trajet car les sources et les récepteurs mis en jeu dans chacune des deux mesures ne sont pas les mêmes. En déplaçant encore deux fois les positions des lignes de sources et de récepteurs, on utilise aussi des ondes planes se propageant du haut vers le bas et du bas vers le haut (voir la figure 2.12b). On obtient ainsi 4 jeux de mesures indépendants des temps de parcours dans la cellule élémentaire. La couverture de cette cellule en terme de densité de trajets est homogène.

4. à partir des temps de parcours mesurés pour les 4 ondes planes considérées, une inversion est réalisée dans la cellule élémentaire afin d'obtenir un modèle de vitesse de groupe des ondes de Rayleigh dans cette cellule. L'algorithme utilisé est celui décrit au paragraphe précédent.
5. la cellule élémentaire est déplacée d'un élément vers la droite et les étapes 1 à 4 sont répétées. En décalant ainsi cette cellule d'un élément à la fois dans les deux directions du plan, l'ensemble de la zone à étudier est couverte.
6. les modèles de vitesse obtenus pour chacune des cellules élémentaires sont combinés afin d'obtenir un modèle de vitesse pour l'ensemble de la zone.

Le résultat de cette inversion est présenté sur la figure 2.13a, où on peut voir les variations de vitesse de groupe des ondes de Rayleigh dues aux hétérogénéités de surface. Si on considère la bande de fréquence utilisée (10-25 Hz) et les vitesses obtenues, les variations de vitesse correspondent à des hétérogénéités situées à des profondeurs jusqu'à 30 m. Le découpage en cellules élémentaires permet d'augmenter la sensibilité de l'inversion aux hétérogénéités locales. Afin de valider l'algorithme d'inversion, la même mesure est effectuée pour différentes tailles de cellule élémentaire (de 8 à 12 éléments dans les lignes de sources et de récepteurs, correspondant à des cellules carrées de 200 à 300 m de côté). Les données mises en jeu pour chaque taille de cellule sont indépendantes. La figure 2.13 montre les résultats ainsi obtenus. L'image de la zone complète reste stable, excepté pour la cellule de 12 éléments. Pour cette dernière taille de cellule, l'erreur sur la mesure est importante car deux ondes se propagent entre les lignes de source et de récepteurs et

le pointé automatique des temps de propagation se fait tantôt sur la première arrivée et tantôt sur la deuxième. Pour les autres tailles de cellule, les modèles obtenus sont d'autant plus lisses que la taille de la cellule est grande, et le contraste des changements de vitesse diminue, ce qui est une conséquence directe de l'augmentation de la distance des trajets considérés pour l'inversion.

Le résultats de l'inversion sont très sensibles à la qualité du pointé des temps de propagation, et une erreur sur ces mesures peut engendrer des modifications importante du modèle final. Au vu de la méthode employée ici pour mesurer les temps de propagation, les erreurs peuvent venir d'un défaut de planéité de l'onde générée à l'aide de la ligne de sources, en particulier pour les récepteurs situés près des coins de la cellule élémentaire utilisée pour l'inversion. Cette mesure pourrait être améliorée en considérant des lignes de sources plus grandes que les lignes de récepteurs afin d'améliorer la qualité de l'onde plane enregistrée.

Une deuxième source d'erreur est la mesure en vitesse de groupe via l'enveloppe du signal. En effet, la détermination du temps du maximum de l'enveloppe est d'autant plus précise que la bande de fréquence considérée est large. Inversement, le milieu étant dispersif (voir la figure 2.10), la vitesse de groupe varie avec la fréquence et l'utilisation d'une bande de fréquences large introduit une erreur sur le pointé du temps de propagation. Il faut alors trouver un compromis et considérer une bande de fréquences assez large pour obtenir un paquet d'onde n'ayant pas une extension trop grande en temps, tout en gardant une dispersion faible.

Utilisation de fonctions de Green reconstruites par corrélation

Le même traitement est utilisé afin d'obtenir une carte de vitesse de groupe à partir des fonctions de Green reconstruites. Dans ce cas, la cellule élémentaire est constituée de deux lignes de récepteurs. La deuxième étape de l'inversion décrite plus haut est alors modifiée comme suit :

2. les fonctions de corrélations entre toutes les paires possibles de récepteurs mettant en jeu un récepteur de la ligne de gauche et un de la ligne de droite sont calculées, en utilisant toutes les sources disponibles. A partir de ces corrélations, une source est simulée à la place de chaque récepteur de la ligne de gauche, et cette ligne de sources virtuelles est utilisée pour former une onde plane se propageant de gauche à droite dans la cellule.

Dans la troisième étape, les ondes planes se propageant de gauche à droite et de droite à gauche (ou de haut en bas et de bas en haut) mettent en jeu exactement les mêmes fonctions de corrélation, mais combinées de façon différentes. Les 4 mesures correspondant aux 4 ondes planes peuvent donc être considérées comme indépendantes.

Le modèle de vitesse obtenu après combinaison des résultats des inversions pour chaque cellule élémentaire est celui de la figure 2.14b. Les deux modèles de vitesse présentés sur cette figure (actif et passif) sont très proches. Les principales différences sont situées sur les bords de la zone étudiée, là où la technique de corrélation est peu performante à cause d'une distribution de sources défavorable. En effet, pour une paire de récepteurs située en bordure du réseau, la distribution de sources n'est pas isotrope et la convergence de la corrélation vers la fonction de Green n'est pas assurée.

Un des avantages de la technique de corrélation est qu'elle nécessite théoriquement beaucoup moins de sources que les techniques actives. En effet, une distribution de sources limitée à la pé-

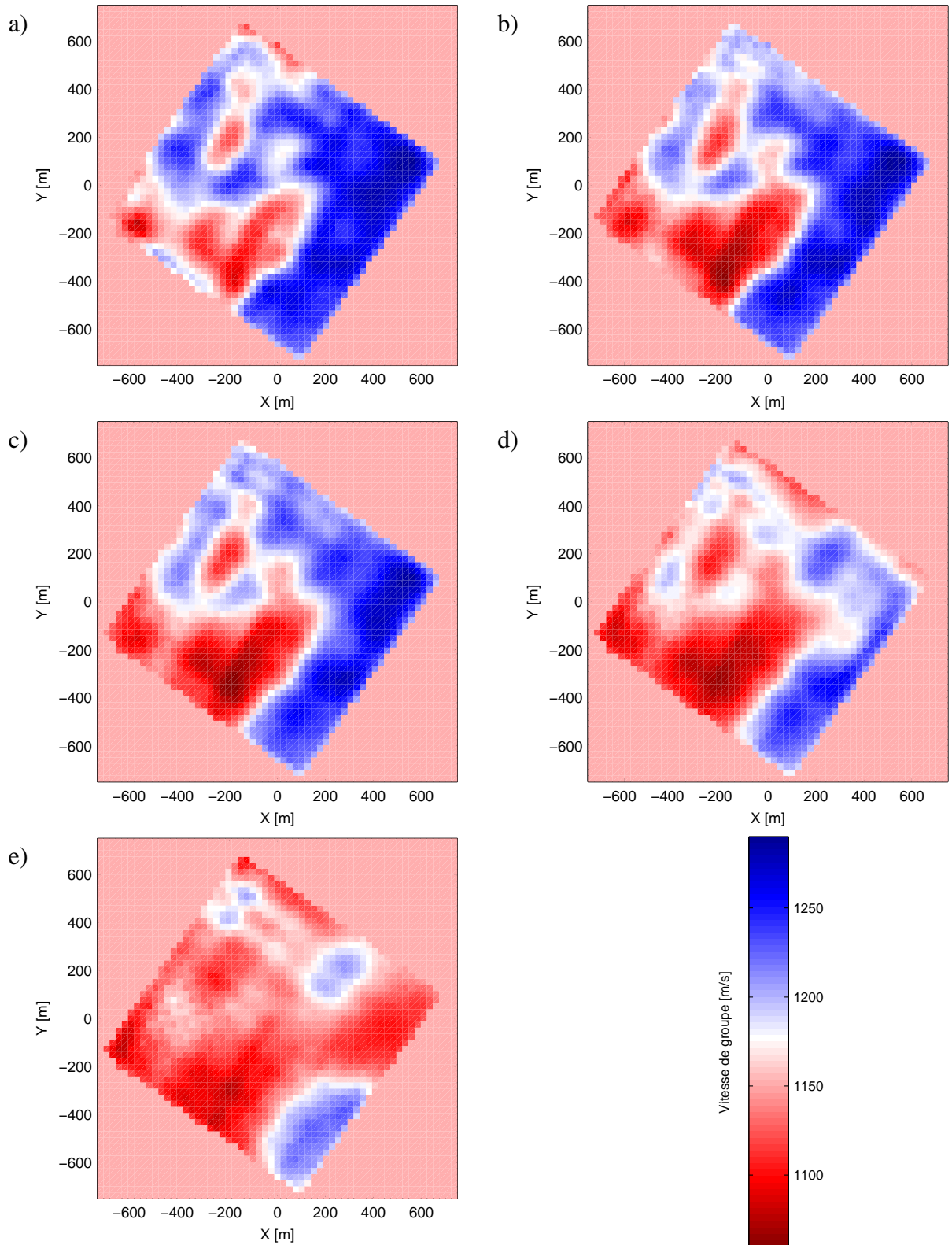


FIG. 2.13 – L’algorithme d’inversion est utilisé pour des tailles cellulaires élémentaires de 8 (figure a.) à 12 (figure e.) éléments dans les lignes de sources et de récepteurs (ces éléments étant séparés de 25 m). L’image est la même quelque soit la taille de cellule utilisée, mis à part le contraste (*i.e.*, la valeur de la vitesse), le modèle obtenu étant d’autant plus lisse que la cellule est grande. Le modèle obtenu avec une cellule de 12 éléments est différent des autres modèles, à cause d’erreurs dans la mesure des temps de propagation.

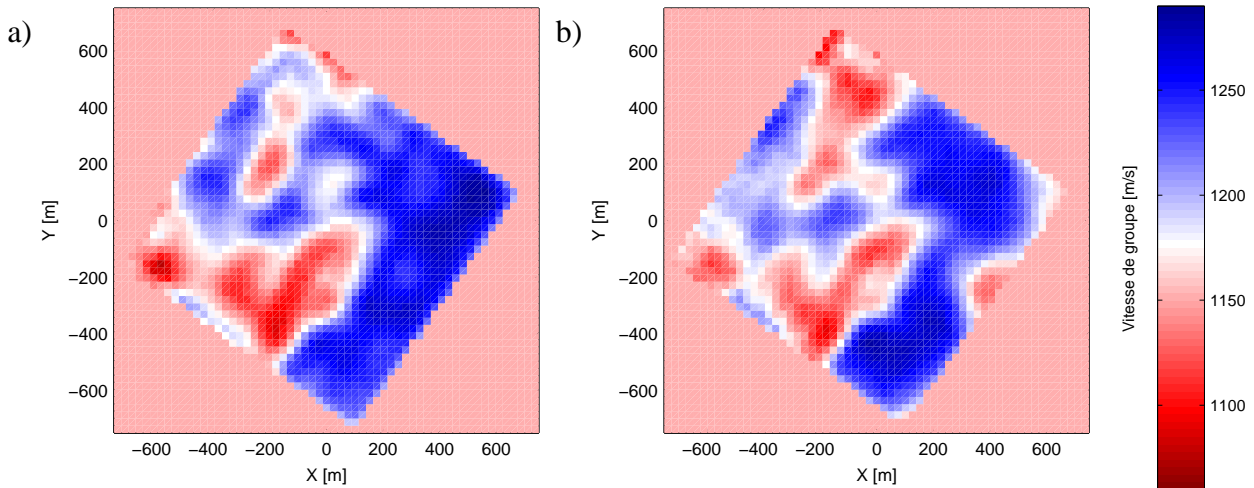


FIG. 2.14 – **a.** Modèle de vitesse de groupe des ondes de Rayleigh obtenu à partir de l'inversion appliquée aux données brutes, pour la zone d'1 km carré étudiée (figure identique à 2.13a). **b.** Le résultat du même algorithme appliqué aux fonctions de Green reconstruites par corrélation. Les deux images ont la même échelle de couleurs. Pour les deux inversions, la taille de la cellule est fixée à 8 éléments.

riphérie de la zone à étudier, en l'entourant complètement, devrait suffire à reconstruire toutes les fonctions de Green nécessaires à l'inversion.

Analyse fréquentielle

La suite de ce travail est présentée en utilisant les données non corrélées. Il serait intéressant d'effectuer le même traitement à partir des données corrélées, mais je n'ai malheureusement pas eu le temps de le faire dans le cadre de ma thèse.

Les cartes de vitesse présentées jusqu'ici ont été obtenues en utilisant les données filtrées entre 10 et 25 Hz. Afin de pouvoir réaliser une tomographie et de retrouver le modèle de vitesse avec la profondeur, il faut pouvoir mesurer les vitesses de groupe à différentes fréquences. J'ai donc repris le traitement à partir de données filtrées dans différentes bandes de fréquences de 5 Hz de largeur pour des fréquences centrales variant de 11,5 à 23,5 Hz par pas de 1 Hz. Pour chaque fréquence, j'ai choisi comme modèle initial un modèle homogène ayant pour vitesse de groupe celle du modèle homogène présenté sur la figure 2.10d pour la fréquence centrale considérée. La variation relative maximale de vitesse est fixée à $\pm 10\%$. Les résultats de ces inversions sont présentés sur la figure 2.15.

Afin de valider l'approche menant à la figure 2.15, j'ai extrait de cette figure des courbes de dispersion pour les 6 positions indiquées par des numéros. Pour chacune de ces positions, j'ai sélectionné un réseau linéaire de 6 récepteurs, ainsi qu'une source située à 200 m, dans son alignement. J'ai alors utilisé ce dispositif pour mesurer les courbes de dispersion en vitesse de groupe de façon classique en utilisant une méthode de *slantstack* appliquée aux enveloppes des sismogrammes, en supposant le milieu localement homogène sous le dispositif. Cette mesure sera considérée comme non biaisée afin d'évaluer la précision de l'approche exposée précédemment. La figure 2.16 montre la comparaison des deux mesures. Elle montre que les modèles de vitesse de groupe présentés sur

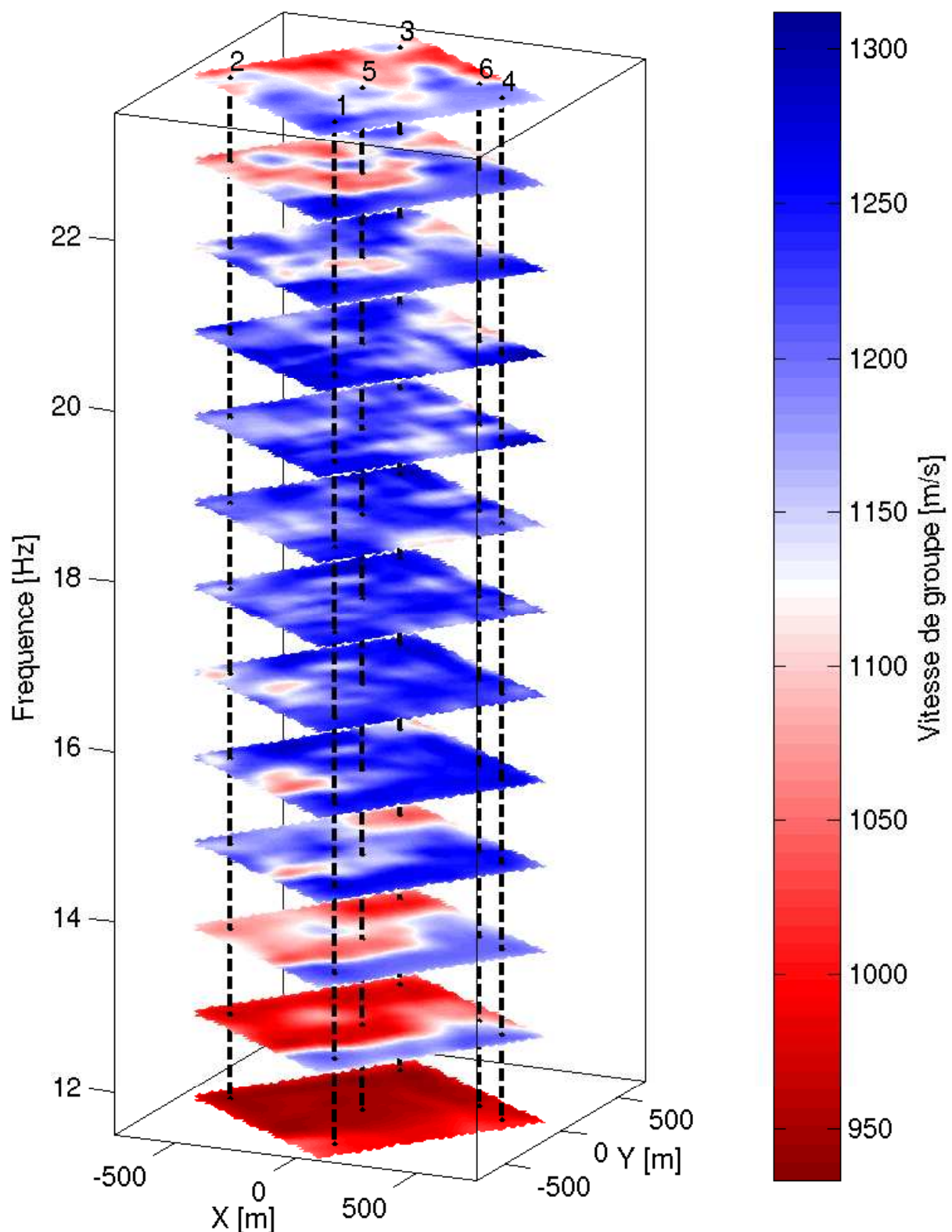
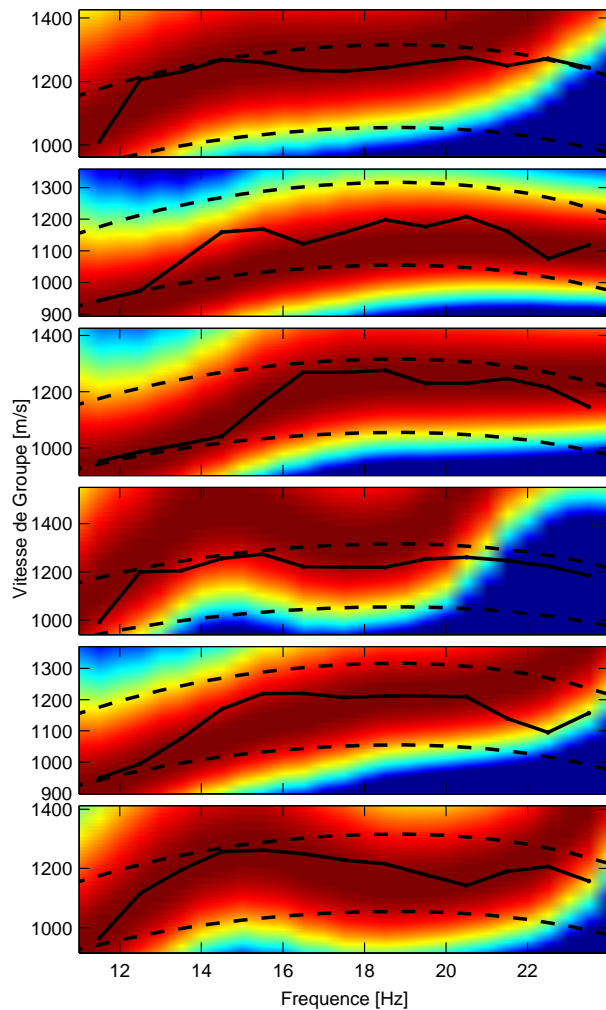


FIG. 2.15 – L'inversion peut être réalisée dans différentes bandes de fréquence. Chaque bande de fréquence fait 5 Hz de large, et elles sont décalée de 1 Hz entre 9 et 21 Hz pour la fréquence basse (soit 9-14 Hz ; 10-15 Hz ; ... ; 21-26 Hz). Les lignes en pointillés et les numéros indiquent les positions où ont été extraites les courbes de dispersion de la figure 2.16.

FIG. 2.16 – Pour chaque position indiquée par les numéros sur la figure 2.15, les courbes de dispersion en vitesse de groupe sont extraites et présentées en trait noir continu (de haut en bas pour les numéros de 1 à 6). Pour chacune de ces positions, une mesure classique des courbes de dispersion (*slantstack* sur les enveloppes des traces) est également réalisée à l'aide d'un réseau de 6 récepteurs, et présentée en niveau de couleurs. Les courbes en pointillés indiquent les limites de variation de vitesse lors de l'inversion utilisée pour obtenir la figure 2.15.



la figure 2.15 ne reproduisent pas toujours les valeurs absolues des vitesses, peut être en partie à cause d'un mauvais choix de vitesse initiale et de variation maximale autorisée dans le processus d'inversion. Les variations de vitesse en fonction de la fréquence sont cependant représentatives des variations réelles, et les courbes présentées sur la figure 2.15 reproduisent les mesures en *slantstack*.

Afin de réaliser une tomographie de la zone menant à un modèle de vitesse des ondes S dans le milieu, en fonction de la profondeur, il faudrait, pour chaque position du modèle de vitesse de groupe des ondes de Rayleigh présenté sur la figure 2.15, inverser la courbe de dispersion. La combinaison de tous les modèles ainsi obtenus permet alors d'obtenir un modèle 3D (ou pseudo-3D) de la zone étudiée.

► 2.4.3 Conclusion et perspectives sur l'imagerie passive

Ces résultats démontrent la faisabilité de l'imagerie de proche surface à partir de fonctions de Green reconstruites par corrélation, aussi bien dans le cas d'un milieu homogène que dans le cas plus complexe d'un milieu latéralement hétérogène. Les résultats obtenus en utilisant des sources virtuelles sont comparables à ceux obtenus à partir des données brutes, parfois même avec une meilleure résolution comme je l'ai montré dans le cas du milieu homogène. Un avantage important

des méthodes de corrélation est qu'elles nécessitent théoriquement moins de sources. Le travail doit être poursuivi dans cette direction afin de mieux quantifier les dégradations de la qualité de l'image finale du milieu lorsqu'on diminue effectivement le nombre de sources.

Le travail d'inversion en vue d'une tomographie a été réalisé en utilisant les données non corrélées. J'ai montré qu'en utilisant les fonctions de Green reconstruites on pouvait obtenir la même image du milieu, mais cette démarche doit être validée. Ici encore, une étude de la dégradation des résultats en fonction du nombre de sources utilisées pour la corrélation pourrait être effectuée. Je me suis arrêté à la mesure des courbes de dispersion en vitesse de groupe en fonction de la position. L'image du milieu présenté sur la figure 2.15 pourrait être à son tour inversée afin d'obtenir une image en vitesse des ondes S en fonction de la profondeur. Cette approche a déjà été utilisée à l'échelle régionale [Shapiro et al., 2005], mais elle reste à être validée pour les échelles utilisées ici.

2.5 Caractérisation de la diffusion : mesure du libre parcours moyen élastique

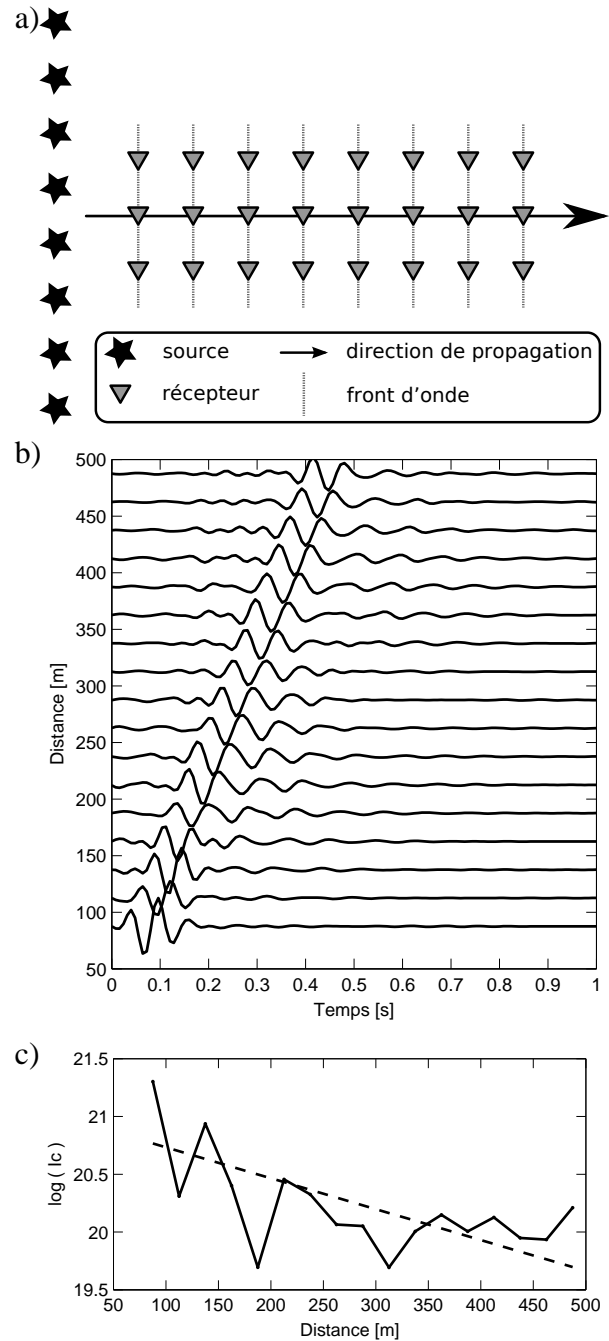
Le milieu étudié ici est karstique et donc diffusant. J'ai voulu caractériser cette diffusion, en effectuant une mesure du libre parcours moyen élastique l . Le libre parcours moyen élastique est la distance caractéristique parcourue par l'onde entre deux événements de diffusion. C'est donc une grandeur qui caractérise l'importance de la diffusion dans le milieu.

J'ai testé deux méthodes de mesure du libre parcours moyen. La première est basée sur la mesure de la décroissance de l'amplitude d'une onde incidente plane dans un milieu diffusant. Si on suppose le milieu sans atténuation (ce qui est le cas ici, voir la figure 2.7), l'amplitude d'une telle onde moyennée sur plusieurs réalisations du milieu, décroît avec la distance parcourue, suivant une loi exponentielle décroissante dont la longueur caractéristique est le libre parcours moyen [Derode et al., 2001]. J'ai donc considéré une ligne de 20 sources s'étendant sur 500 m. Cette ligne est utilisée pour former une onde plane, que l'on enregistre le long d'une ligne de récepteurs, perpendiculaire et centrée par rapport à la ligne de sources (figure 2.17a). Afin d'améliorer la mesure de l'onde plane, chaque point de mesure est constitué d'un groupe de 5 récepteurs alignés et perpendiculaires à la direction de propagation, et les traces enregistrées sur chacun des 5 récepteurs à chaque distance sont sommées (cette technique, appelée formation de voie, est décrite plus en détails au paragraphe 5.1.1). La moyenne sur les réalisations est ici remplacée par un déplacement du réseau de sources et de récepteurs (comme décrit dans [Derode et al., 2001]). 100 réalisations sont ainsi considérées, et les 100 sections sismiques obtenues sont moyennées pour obtenir celle présentée sur la figure 2.17b. Dans le cas d'un milieu diffusant non réverbérant et non atténuant, l'amplitude de cette onde plane décroît selon une exponentielle dont la longueur caractéristique est liée au libre parcours moyen élastique :

$$\langle s(t, r) \rangle = \exp\left(-\frac{r}{2l}\right) e^{-i(\omega t - \vec{k} \cdot \vec{r})} , \quad (2.15)$$

où $\langle s(t, r) \rangle$ est le signal enregistré en fonction du temps et de la distance, moyenné de façon cohérente sur les réalisations, l est le libre parcours moyen, ω la pulsation et \vec{k} le vecteur d'onde réel

FIG. 2.17 – **a.** Schéma du réseau utilisé pour la mesure de la décroissance de l'amplitude d'une onde plane incidente. Les lignes de sources et de récepteurs sont représentées ici avec seulement 8 éléments, alors qu'elles en comportent 20 pour la mesure. Une onde plane est formée à l'aide de la ligne de source. Cette onde plane est enregistrée le long de 5 ligne de récepteurs parallèles (3 seulement sont représentées ici). Les enregistrements sur ces 5 lignes sont sommés pour former $s(t, r)$ utilisé dans l'équation (2.15). Le réseau est ensuite déplacé afin d'obtenir 100 réalisations pour le milieu. **b.** Section sismique obtenue en moyennant des ondes planes enregistrées en 100 positions différentes. **c.** Évolution du logarithme naturel de l'amplitude (maximum de l'enveloppe) de l'onde plane en fonction de la distance. La ligne en pointillés indique la décroissance attendue pour le libre parcours moyen mesuré grâce au rapport entre les intensités cohérentes et incohérentes, soit 374 m.



dans le milieu diffusant. La formation d'une onde plane est nécessaire afin de ne pas introduire de dispersion géométrique dans la décroissance du champ. Afin d'évaluer l'amplitude de l'onde plane sur la section sismique de la figure 2.17b, j'ai mesuré le maximum de l'enveloppe du signal à chaque distance, dans la bande 10-15 Hz. Le logarithme naturel de cette amplitude est tracé sur la figure 2.17c. Selon l'équation (2.15), ce logarithme devrait évoluer linéairement avec la distance, ce qui n'est pas le cas. On remarque en particulier que l'amplitude diminue fortement aux alentours de 150 m, avant de remonter. Ces fluctuations peuvent être dues à des interférences entre deux ondes. Les hypothèses permettant d'établir l'équation (2.15) sont donc mises en défaut : le milieu semble être réverbérant.

J'ai donc cherché à mesurer le libre parcours moyen par une autre méthode prenant en compte ces deux cas. Le principe de cette méthode est de normaliser l'amplitude du champ moyen cohérent par celle du champ moyen incohérent. La notion de moyenne cohérente ou incohérente sur les réalisations est ici introduite. La moyenne du champ est dite cohérente lorsqu'elle prend en compte la phase des signaux. Des interférences entre les champs d'onde ainsi sommés sont donc possibles. Par opposition, une moyenne est incohérente si elle ne tient pas compte de la phase des signaux. La moyenne incohérente considérée ici sera effectuée sur l'intensité du champ.

Les détails de la méthode sont décrits dans [de Rosny & Roux, 2001] et [de Rosny & Roux, 2004] pour une cavité réverbérante. Je vais exposer ici son application au jeu de données Qarn Alam, pour lequel on suppose un milieu avec réverbération. On considère le signal $s_n(t, \vec{r})$ issu d'une source n et enregistrée sur le réseau. On peut alors former la moyenne cohérente des signaux enregistrés à la même distance de la source : $s_n(t, r) = \langle s_n(t, \vec{r}) \rangle_{\|\vec{r}\|=r}$. Cette première moyenne sur les récepteurs est une moyenne sur des réalisations différentes du milieu puisque les trajets source-récepteurs considérés sont différents. L'étape suivante consiste à moyennner ce signal pour l'ensemble des sources du réseau : $s(t, r) = \langle s_n(t, r) \rangle_n$. On calcule enfin l'intensité cohérente moyennée sur tous les enregistrements comme :

$$I_c(t, r) = s^2(t, r) = \left\langle \langle s_n(t, \vec{r}) \rangle_{\|\vec{r}\|=r} \right\rangle_n^2. \quad (2.16)$$

Une deuxième moyenne d'ensemble est calculée afin de former l'intensité incohérente :

$$I_i(t, r) = \left\langle \langle s_n^2(t, \vec{r}) \rangle_{\|\vec{r}\|=r} \right\rangle_n. \quad (2.17)$$

La dépendance de l'amplitude du champ d'onde cohérent $s(t, r)$ avec le libre parcours moyen peut être décrit selon une équation similaire à l'équation (2.15). A cette décroissance due à la diffusion viennent s'ajouter les contributions de l'atténuation dans le milieu et de la dispersion géométrique, les ondes considérées ici n'étant pas planes. Si de plus le milieu est réverbérant, les ondes issues de réflexions sur des interfaces modifient l'amplitude du champ enregistré, celui étant maintenant composé de plusieurs ondes. De tels phénomènes peuvent expliquer l'évolution de l'amplitude du champ avec la distance observée sur la figure 2.17c. Afin d'isoler la contribution de la diffusion dans la décroissance de l'amplitude du champ, l'intensité cohérente peut être normalisée par l'intensité incohérente, formant ainsi le rapport $R(t, r) = I_c(t, r) / I_i(t, r)$. L'intensité incohérente étant sensible à l'atténuation, à la réverbération et à la dispersion géométrique, mais

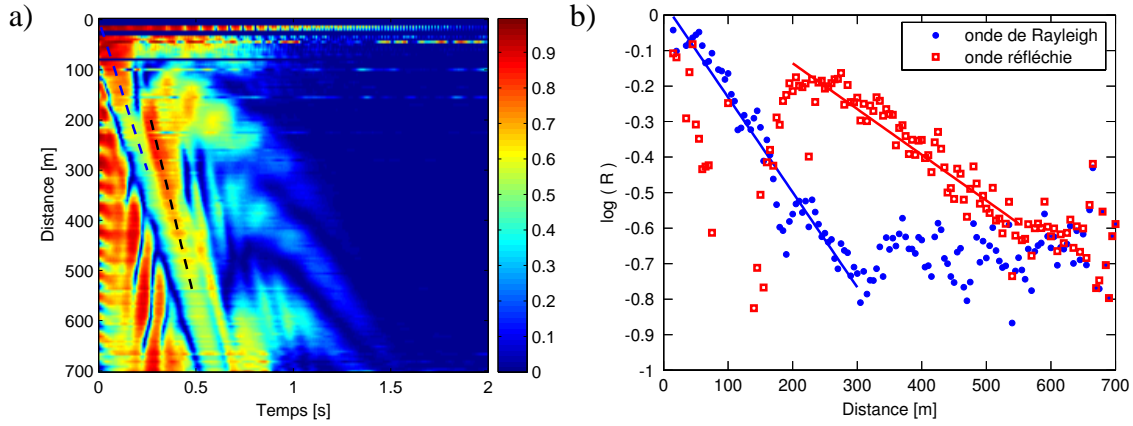


FIG. 2.18 – **a.** Le rapport $R(r,t) = I_c(r,t)/I_i(r,t)$. Deux ondes sont visibles sur cette figure : une onde de Rayleigh lente et une onde réfléchie plus rapide. Les lignes en pointillés indiquent les temps d'arrivée considérés pour les deux ondes, qui correspondent à des vitesses de 1200 et 1600 m/s. Le logarithme naturel des valeurs du rapport R le long de chacune de ces lignes sont tracées en **b.**, ainsi que les régressions linéaires permettant de mesurer les libres parcours moyens élastiques l selon l'équation (2.18).

pas à la diffusion, seule sa contribution reste dans le rapport R . Pour une onde se propageant dans le milieu à une vitesse c , on considère la valeur de R prise au temps de passage de l'onde :

$$R(t) = R(t = r/c, r) = \frac{I_c(r/c, r)}{I_i(r/c, r)} = \exp\left(-\frac{r}{l}\right) \quad (2.18)$$

Dans un premier temps, j'ai calculé le rapport $R(r,t)$, sans le limiter au temps d'arrivée des ondes. Le résultat est présenté sur la figure 2.18a, où deux ondes sont visibles : la plus lente est une onde de Rayleigh et la plus rapide une onde dont le profil est celui d'une onde réfléchie, et que je désignerai donc ainsi, même si le manque de connaissance sur le milieu ne permet pas de confirmer sa nature. La vitesse de propagation de chacune de ces ondes est mesurée sur la figure, ainsi que les plages de distances sur lesquelles elles sont identifiables sans ambiguïté. Les vitesses mesurées sont de 1200 m/s pour l'onde de Rayleigh et 1600 m/s pour l'onde réfléchie, mesurées respectivement dans les intervalles de distances 0 à 300 m et 200 à 550 m. La figure 2.18b montre l'évolution du logarithme naturel du rapport $R(r)$ en fonction de la distance pour chacune de ces deux ondes. Une régression linéaire est utilisée pour trouver les libres parcours moyens dans les deux cas grâce à l'équation (2.18). Les valeurs trouvées sont 374 m pour l'onde de Rayleigh et 776 m pour l'onde réfléchie. La différence entre ces deux valeurs peut être expliquée par la différence de nature entre les deux ondes, la première étant une onde de surface et la deuxième une onde de volume. L'onde réfléchie étant une onde de volume elle est sensible à des profondeurs plus importantes que l'onde de Rayleigh, la diffusion en profondeur étant *a priori* différente de celle en surface. L'accord avec la loi théorique est meilleur ici que pour la mesure de la décroissance de l'amplitude de l'onde plane incidente décrite plus haut. La valeur du libre parcours moyen pour l'onde de Rayleigh est en accord avec la courbe de la figure 2.17b. La présence de deux ondes dans le champ enregistré explique les fluctuations dans la mesure de l'amplitude sur cette figure, dues aux interférences entre ces ondes.

La figure 2.18b montre une saturation du rapport R pour l'onde de Rayleigh au delà d'une certaine distance. Cette saturation est liée au nombre N_r de réalisations indépendantes dans le milieu

[de Rosny & Roux, 2001], la plateau ayant pour valeur $1/N_r$. La valeur de $e^{-0,7}$ obtenue ici donne un nombre de réalisations indépendantes $N_r = 2$. Ce nombre est en accord avec la valeur de libre parcours moyen élastique $l = 374 \text{ m}$ obtenue pour cette onde, comparée à la taille du réseau d'un kilomètre carré.

Bibliographie chapitre 2

Aki, K., 1957. Space and time spectra of stationary stochastic waves with special reference to microtremors, *Bulletin of the Earthquake Research Institute*, **35**, 415–456.

Campillo, M., 2006. Phase and correlation in ‘random’ seismic fields and the reconstruction of the Green function, *Pure and Applied Geophysics*, **163**(2-3), 475–502.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**, 547–549.

Campman, X. H., van Wijk, K., Scales, J. A., & Herman, G. C., 2005. Imaging and suppressing near-receivers scattered surface waves, *Geophysics*, **70**(2), V21–V29.

Claerbout, J. F., 1968. Synthesis of a layered medium from its acoustics transmission response, *Geophysics*, **33**, 264–269.

Colin de Verdière, Y., 2006a. Mathematical models for passive imaging I: general background.

Colin de Verdière, Y., 2006b. Mathematical models for passive imaging II: effective hamiltonians associated to surface waves.

de Rosny, J. & Roux, P., 2001. Multiple scattering in a reflecting cavity: Application to fish counting in a tank, *The Journal of the Acoustical Society of America*, **109**(6), 2587–2597.

de Rosny, J. & Roux, P., 2004. Reply to the comment on "Multiple scattering in a reflecting cavity: Application to fish scattering" [J. Acoust. Soc. Am. 113, 2978-2979 (2003)] (1), *The Journal of the Acoustical Society of America*, **115**(1), 31–34.

Derode, A., Tourin, A., & Fink, M., 2001. Random multiple scattering of ultrasound. I. Coherent and ballistic waves, *Physical Review E*, **64**(3), 036605.

Derode, A., Larose, E., Campillo, M., & Fink, M., 2003a. How to estimate the Green’s function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Applied Physics Letters*, **83**(15), 3054–3056.

Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M., & Fink, M., 2003b. Recovering the Green’s function from field-field correlations in an open scattering medium (L), *The Journal of the Acoustical Society of America*, **113**(6), 2973–2976.

Draganov, D., Wapenaar, K., Mulder, W., Singer, J., & Verdel, A., 2007. Retrieval of reflections from seismic background-noise measurements, *Geophysical Research Letters*, **34**, L04305, - ondes reflechies.

Duvall, T. L., Jefferies, S. M., Harvey, J. W., & Pomerantz, M. A., 1993. Time distance helioseismology, *Nature*, **362**, 430–432.

Gilles, P. M., Duvall, T. L., Scherrer, P. H., & Bogart, R. S., 1997. A subsurface flow of material from the Sun’s equator to its poles, *Nature*, **390**, 52–54.

- Godin, O. A., 2007. Emergence of the acoustic Green's function from thermal noise, *The Journal of the Acoustical Society of America*, **121**(2), EL96–EL102.
- Gouédard, P., Roux, P., Campillo, M., & Verdel, A., 2008a. Convergence of the two-point correlation function toward the green's function in the context of a seismic prospecting dataset, *Geophysics*, p. in press.
- Gouédard, P., Stehly, L., Brenguier, F., Campillo, M., Colin de Verdière, Y., Larose, E., Margerin, L., Roux, P., Sánchez-Sesma, F. J., Shapiro, N. M., & Weaver, R. L., 2008b. Cross-correlation of random fields: mathematical approach and applications, *Geophysical Prospecting*, **56**(3), 375–393.
- Hennino, R., Tégourès, N., Shapiro, N. M., Margerin, L., Campillo, M., van Tiggelen, B. A., & Weaver, R. L., 2001. Observation of equipartition of seismic waves, *Physical Review Letters*, **86**(15), 3447–3450.
- Herman, G. C. & Perkins, C., 2006. Predictive removal of scattered noise, *Geophysics*, **71**(2), V41–V49.
- Larose, E., Derode, A., Campillo, M., & Fink, M., 2004. Imaging from one-bit correlation of wide-band diffuse wavefield, *Journal of Applied Physics*, **95**(12), 8393–8399.
- Larose, E., Margerin, L., Derode, A., van Tiggelen, B. A., Campillo, M., Shapiro, N. M., Paul, A., Stehly, L., & Tanter, M., 2006. Correlation of random wavefields: an interdisciplinary review, *Geophysics*, **71**(4), SI11–SI21.
- Larose, E., Roux, P., & Campillo, M., 2007. Reconstruction of rayleigh-lamb dispersion spectrum based on noise obtained from an air-jet forcing, *The Journal of the Acoustical Society of America*, **122**(6), 3437–3444.
- Lobkis, O. I. & Weaver, R. L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *The Journal of the Acoustical Society of America*, **110**(6), 3011–3017.
- Paul, A. & Campillo, M., 2001. Extracting the Green function between two stations from coda waves, in *AGU Fall Meeting Abstracts*, p. D610.
- Rickett, J. & Claerbout, J. F., 1999. Acoustic daylight imaging via spectral factorization; helioseismology and reservoir monitoring, *The Leading Edge*, **18**(8), 957–960.
- Roux, P. & Kuperman, W. A., 2004. Extracting coherent wavefronts from acoustic ambient noise in the ocean, *The Journal of the Acoustical Society of America*, **116**(4), 1995–2003.
- Roux, P., Sabra, K. G., Kuperman, W. A., & Roux, A., 2005. Ambient noise cross correlation in free space: theoretical approach, *The Journal of the Acoustical Society of America*, **117**(1), 79–84.
- Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C., 2005a. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophysical Research Letters*, **32**, L03310.

- Sabra, K. G., Roux, P., & Kuperman, W. A., 2005b. Emergence rate of the time-domain Green's function from the ambient noise cross-correlation, *The Journal of the Acoustical Society of America*, **118**(6), 3524–3531.
- Sabra, K. G., Srivastava, A., di Scalea, F. L., Bartoli, I., Rizzo, P., & Conti, S., 2008. Structural health monitoring by extraction of coherent guided waves from diffuse fields, *The Journal of the Acoustical Society of America*, **123**(1), EL8–EL13.
- Sánchez-Sesma, F. J. & Campillo, M., 2006. Retrieval of the Green function from cross correlation: the canonical elastic problem, *Bulletin of the Seismological Society of America*, **96**(3), 1182–1191.
- Sánchez-Sesma, F. J., Pérez-Ruiz, J. A., Luzón, F., Campillo, M., & Rodríguez-Castellanos, A., 2008. Diffuse fields in dynamic elasticity, *Wave Motion*, **45**(5), 641–654.
- Schuster, G. T., Yu, J., Sheng, J., & Rickett, J., 2004. Interferometric/daylight seismic imaging, *Geophysical Journal International*, **157**(2), 838–852.
- Shapiro, N. M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophysical Research Letters*, **31**, L07614.
- Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Physical Review E*, **69**(4), 046610.
- Tarantola, A., 1987. *Inverse problem theory: methods for data fitting and model parameter estimation*, Amsterdam.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Reviews of Geophysics and Space Physics*, **20**, 219–232.
- van Tiggelen, B. A., 2003. Green function retrieval and time reversal in a disordered world, *Physical Review Letters*, **91**(24), 243904.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's Function of an arbitrary inhomogeneous medium by cross-correlation, *Physical Review Letters*, **93**(25), 254301.
- Wapenaar, K., 2006. Nonreciprocal Green's function retrieval by cross correlation, *The Journal of the Acoustical Society of America*, **120**(1), EL7–EL13.
- Wapenaar, K., Fokkema, J., & Snieder, R., 2005. Retrieving the Green's function by cross-correlation: a comparison of approaches (I), *The Journal of the Acoustical Society of America*, **118**(5), 2783–2786.
- Weaver, R. L. & Lobkis, O. I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Physical Review Letters*, **87**(13), 134301.

Weaver, R. L. & Lobkis, O. I., 2003. Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal phonons, *The Journal of the Acoustical Society of America*, **113**(5), 2611–2621.

Mesure de courbes de dispersion à partir de corrélations de bruit

Sommaire

3.1 Méthodes passives d'analyse en réseau	82
3.1.1 La méthode <i>Spatial Autocorrelation</i> (SPAC)	82
3.1.2 La méthode fréquence/nombre d'onde (FK)	82
3.2 Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique	83
3.2.1 Summary	84
3.2.2 Introduction	84
3.2.3 Ambient noise simulation and array configuration	86
3.2.4 Phase velocity dispersion curves obtained from HRFK and SPAC analysis	90
3.2.5 Phase velocity dispersion curves obtained from noise correlations slantstack (NCSS)	92
3.2.6 Results comparison	94
3.2.7 Conclusion	98
Bibliographie chapitre 3	100

Nous avons vu au chapitre précédent que la fonction de Green d'un milieu entre deux points peut être reconstruite par corrélation. Nous allons maintenant voir comment utiliser cette technique pour mesurer des courbes de dispersion à partir de bruit de fond. Afin d'évaluer la précision de la mesure, j'ai utilisé des données synthétiques calculées pour un milieu multi-couches complexe mais parfaitement connu. Différentes stratégies sont présentées en fonction de la distribution des sources de bruit, dont l'importance a été mise en évidence dans le chapitre 2. L'erreur de mesure de la courbe de dispersion est comparée à celle faite en utilisant les méthodes passives d'analyse en réseau, classiquement utilisées en géophysique.

Le principe des méthodes d'analyse en réseau est d'abord présenté, avant d'aborder la comparaison avec la technique de corrélation. Ce dernier point a fait l'objet d'une publication [Gouédard et al., 2008].

3.1 Méthodes passives d'analyse en réseau

Les méthodes d'analyse en réseaux sont apparues vers la fin des années 50, avec les travaux d'Aki et la méthode *Spatial Autocorrelation* [méthode SPAC, Aki, 1957, 1965]. Un peu plus tard, une autre famille de méthodes en réseau opérant dans les espaces réciproques est apparue (méthodes en fréquence/nombre d'onde ou FK), avec les travaux de Capon [1969] et Lacoss et al. [1969]. Depuis, ces méthodes n'ont cessé d'être développées. Nous allons présenter ici les deux méthodes les plus courantes permettant la mesure de courbes de dispersion à partir du bruit de fond enregistré sur un réseau.

► 3.1.1 La méthode *Spatial Autocorrelation* (SPAC)

L'idée d'Aki [1957] était qu'il n'est pas nécessaire de connaître la source du bruit ambiant pour mesurer les temps de propagation entre les éléments d'un réseau. Il a alors défini la fonction d'autocorrélation spatiale entre deux récepteurs A et B séparés par une distance R par

$$C(R) = \frac{1}{T} \int_0^T s_A(t) s_B(t) dt \quad , \quad (3.1)$$

où s_A et s_B sont les signaux enregistrés pendant un temps T en A et B . Si les signaux sont filtrés dans une bande étroite autour de la pulsation ω , on peut définir le rapport d'autocorrélation par

$$\rho(R, \omega) = \frac{C(R, \omega)}{C(0, \omega)} \quad . \quad (3.2)$$

Ce rapport est calculé pour toutes les paires d'un réseau de récepteurs. Pour une distance donnée R , Aki [1957] a montré que la moyenne azimutale de $\rho(R, \omega)$ était proportionnelle à une fonction de Bessel d'ordre zéro :

$$\rho(R, \omega) = J_0 \left(\frac{\omega R}{c(\omega)} \right) = J_0(Rk) \quad , \quad (3.3)$$

où $c(\omega)$ est la vitesse de phase des ondes dans le milieu, pour la pulsation ω (*i.e.*, la courbe de dispersion), et k le nombre d'onde.

Cette méthode repose sur deux hypothèses : 1) les sources de bruit sont réparties de façon aléatoires ; et 2) il n'y a qu'une vitesse de phase par fréquence (*i.e.*, il n'y a qu'un seul mode de propagation). La première de ces hypothèses est rarement vérifiée. Elle reste cependant acceptable tant que le bruit sismique n'est pas très directif, et cette méthode est performante dans la pratique tant que les sources de bruit sont nombreuses et localisées tout autour du réseau.

► 3.1.2 La méthode fréquence/nombre d'onde (FK)

Cette méthode, développée par Capon [1969] et Lacoss et al. [1969], opère dans les espaces réciproques. Je l'ai déjà décrite dans ce manuscrit au paragraphe 2.3.1 dans le cas particulier des réseaux linéaires. Je traiterai ici le cas des réseaux ayant une extension spatiale non nulle dans les deux directions du plan.

La transformée de Fourier des signaux est calculée, par rapport au temps et à l'espace. La méthode suppose ensuite que des ondes planes horizontales se propagent à travers le réseau. Si on considère alors une telle onde de fréquence f , provenant d'une direction donnée, et se propageant à une vitesse c , on peut calculer les temps d'arrivée relatifs de cette onde sur chaque capteur du réseau. On modifie alors la phase de chacun des signaux en accord avec les temps d'arrivée relatifs. La réponse du réseau pour l'onde considérée est alors définie comme l'énergie de la somme des signaux déphasés. Si l'onde considérée était effectivement présente dans les enregistrements, les déphasages introduits permettent de remettre les signaux en phase sur chacun des récepteurs du réseau, et l'énergie de la somme des signaux est élevée. Inversement, si l'onde n'était pas présente, les signaux se somment de façon incohérente, et l'énergie de la somme est faible. Le résultat est généralement présenté dans un plan k_x/k_y (les deux composantes du vecteur d'onde \vec{k} de l'onde considérée) pour chaque fréquence. Les maximums locaux de la réponse du réseau dans ce plan permettent de retrouver les ondes (*i.e.*, leur direction et leur vitesse) effectivement présentes dans les enregistrements, pour la fréquence considérée.

Une amélioration de cette méthode, appelée *High-Resolution Frequency-Wavenumber* (HRFK, aussi appelée *Minimum Variance Distortionless Response*), a été introduite par Capon [1969]. Il proposa de rajouter une pondération pour chaque capteur dans la somme sur les signaux déphasés. Cette pondération est calculée via une optimisation visant à minimiser la réponse du réseau dès qu'on s'éloigne du vecteur d'onde considéré. Autrement dit, les coefficients de pondération permettent de réduire la taille du pic de réponse dans le plan k_x/k_y . Ces coefficients dépendent de la fréquence et de la position des récepteurs. Ces méthodes, basées sur la comparaison avec des synthétiques, sont performantes tant que le nombre de sources de bruit est limité. En effet, lorsqu'on augmente le nombre de sources de bruit, on peut avoir des interférences entre les ondes produites par chacune d'elles, ce qui n'est pas pris en compte par la méthode.

Les méthodes FK utilisent exactement les mêmes algorithmes que les méthodes de formation de voies décrites au paragraphe 5.1.1. L'approche est néanmoins différente, ces dernières ayant pour but de localiser les sources de bruit en supposant le modèle de vitesse connu. Le lecteur cherchant plus de détails sur les méthodes FK, et en particulier sur l'optimisation mise en jeu dans la méthode HRFK, est invité à se référer au chapitre 5.

3.2 Phase–velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique

Cette section est une retranscription de l'article [Gouédard, Cornou, & Roux, 2008] publié au cours de ma thèse dans la revue *Geophysical Journal International*. Dans cet article nous avons voulu vérifier de façon numérique que la corrélation de bruit permet bien de retrouver les courbes de dispersion dans un milieu multi-couches complexe où plusieurs modes se propagent, et d'en évaluer la précision. Nous avons de plus comparé les résultats de la méthode « NCSS » (pour *Noise Correlation Slant Stack*) à ceux obtenus en appliquant les techniques SPAC [Aki, 1957] et HRFK [Capon, 1969], classiquement utilisées dans les études géotechniques passives. Cette comparaison a

été réalisée dans le cas de bruits isotrope et directionnel. Elle montre que la méthode NCSS permet une mesure des courbes de dispersion à plus haute fréquence que les méthodes classiques, et avec une très grande précision lorsque les propriétés du champ d'onde sont favorables.

► 3.2.1 Summary

It has been demonstrated both theoretically and experimentally that the Green's function between two receivers can be retrieved from the cross-correlation of isotropic noise records. Since surface waves dominate noise records in geophysics, tomographic inversion using noise correlation techniques have been performed from Rayleigh waves so far. However, very few numerical studies implying surface waves have been conducted to confirm the extraction of the true dispersion curves from noise correlation in a complicated soil structure. In this paper, synthetic noise has been generated in a small-scale (<1 km) numerical realistic environment and classical processing techniques are applied to retrieve the phase velocity dispersion curves, first step toward an inversion. We compare results obtained from spatial autocorrelation method (SPAC), high-resolution frequency-wavenumber method (HRFK) and noise correlation slantstack techniques on a 10-sensor array. Two cases are presented in the (1–20 Hz) frequency band that corresponds to an isotropic or a directional noise wavefield. Results show that noise correlation slantstack provides very accurate phase velocity estimates of Rayleigh waves within a wider frequency band than classical techniques and is also suitable for accurately retrieving Love waves dispersion curves.

Key words: Surface waves and free oscillations; Computational seismology.

► 3.2.2 Introduction

The Green's function of a medium between two points A and B represents the record we would get at A if we put an impulse source at B. The use of random noise to reconstruct the Green's function has already been applied successfully in various fields of wave physics such as helioseismology, oceanography or geophysics at large scales and ultrasonics or non-destructive evaluation at small scales. Historically speaking, helioseismology was the first field where ambient noise cross-correlation performed from recordings of the Sun surface random motion was used to produce coherent images of the sun [Duvall et al., 1993, Gilles et al., 1997]. More recently, a seminal paper [Weaver & Lobkis, 2001] has shown how diffuse thermal noise recorded and cross-correlated at two ultrasonic transducers fastened to one face of a duralumin sample provided the complete Green's function between these two points. Similar results were then obtained at a much larger scale in shallow underwater acoustics where both direct and reflected wave fronts were retrieved from ambient noise cross-correlation [Roux & Kuperman, 2004]. At the same time, the cross-correlation process was investigated from the perspective of multiple uncorrelated sources that were considered as noise sources [Campillo & Paul, 2003, Derode et al., 2003]. By summing the contribution of all sources to the correlation, it has been shown numerically that the correlation contains the causal and acausal Green's function of the medium [Wapenaar, 2004]. At ultrasonic scales, the noise-extracted Green's function has been used to detect defects in duralumin samples [Larose et al., 2006a].

In seismology, Aki [1957] proposed fifty years ago to use seismic noise to retrieve the dispersion properties of the subsoil. For elastic waves, it has been theoretically showed that the conver-

gence of noise correlation to the Green's function was bounded by the equipartition condition of the different components of the elastic field [Sánchez-Sesma & Campillo, 2006, Sánchez-Sesma et al., 2006]. In other words, the emergence of the Green's function is effective after a sufficient self-averaging process that is provided by random spatial distribution of the noise sources when considering long time series as well as scattering [Campillo, 2006, Larose et al., 2006b]. Shapiro & Campillo [2004] reconstructed the surface wave part of the Green's function by correlating seismic noise at stations separated by distance of hundreds of kilometres, and measured their dispersion curves at periods ranging from 5 to 20 s. This method led to the first application of passive seismic imaging in California [Sabra et al., 2005, Shapiro et al., 2005]. Since then, several studies focused on seismic tomography at lithospheric scale by using such noise correlation techniques [Yao et al., 2006, Bensen et al., 2007, Lin et al., 2007, Pedersen et al., 2007]. Some authors have also used this technique to address the issue of noise source spatial distribution in the cross-correlation process [Pedersen et al., 2007, Stehly et al., 2007]. However, despite convincing tomographic images of the subsurface shear velocity at large scales [Sabra et al., 2005, Shapiro et al., 2005] there is no study that confirms that the noise correlation process yields the true phase velocity dispersion curves in a realistic environment where several surface wave modes are present. Such a work would imply (1) the computation of the elastic field produced by multiple noise sources in a given subsurface model, (2) the recording of the numerical noise traces on a seismic array, (3) the computation of the noise-correlation function between each receiver pair and (4) the extraction of the phase-velocity dispersion curves from the noise-correlation process for a comparison to the actual dispersion curves directly obtained from the numerical model. This is the objective of this paper in the frame of earthquake engineering implying small-scale seismic networks (<500 m) and high-frequency seismic noise (>1 Hz).

In earthquake engineering, microseisms and microtremors (noise produced by anthropogenic sources) have been used for more than 30 yr for estimating subsurface S-wave velocities and for site effect purposes. Microtremor studies indeed originated in the pioneering work of Japanese authors [Kanai & Tanaka, 1954, Aki, 1957, Nogoshi & Igarashi, 1971, Nakamura, 1989]. In recent decades, the use of microtremor array measurements (mainly frequency-wavenumber based methods [Burg, 1964, Capon, 1969] and the spatial autocorrelation (SPAC) technique [Aki, 1957]) applied to ambient-vibration noise wavefields have spread throughout the world as a substitute of borehole measures or active seismic methods, which are costly and even prohibitive in urban areas. Recently, new methods have emerged: Bettig et al. [2001], Ohori et al. [2002], Asten et al. [2004], Cho et al. [2004, 2006] proposed modification and/or extension of the SPAC technique, while Louie [2001] presented the linear slantstack method for ambient noise recordings. Although noise correlation technique gave rise to numerous applications and studies in large-scale seismology since Shapiro & Campillo [2004], this technique was only applied by Chávez-García et al. [2005] to subsurface structure imaging.

In summary, the goal of microtremor array measurements and noise correlation techniques is to measure dispersion curves from which subsurface shear wave profiles are then extracted. It has to be noted however that even if microtremor array analysis is performed from 2-D arrays, the interpretation of the resulting dispersion curves are mostly performed assuming wave propagation in 1-D structures.

Within the framework of the last symposium on the Effects of Surface Geology on seismic

motion (ESG2006), an international noise blind test [Cornou et al., 2006] was proposed in order to compare surface wave dispersion characteristics derived from competing analysis approaches and to make a clear assessment regarding the potential of microtremor array studies for site effect estimation. This blind test involved both synthetic and real data sets and was opened to a large scientific community, with no restrictions regarding the choice of analysis approaches. Synthetic data provided the opportunity to perform a benchmark test where the site structure and the wavefield situation are fully controlled. For SPAC and FK based techniques — which were the most frequently used techniques by participants — this exercise pinpointed that surface wave dispersion characteristics were well retrieved over frequency bands controlled by the resolution capabilities of seismic array and the energy content of ambient noise. Interestingly, the only participating group that used noise correlation [Gouédard et al., 2006] has shown very good phase velocities estimates for both Rayleigh and Love fundamental and higher modes outside the classical array proxy-capabilities at high frequency, estimates at low frequency being as good as the ones provided by classical SPAC and FK techniques.

The scope of this paper is twofold. First, we numerically demonstrate that the noise-correlation process leads to the true phase-velocity dispersion curves even when several surface wave modes are present. Second, we clarify reasons why noise correlation enables to provide phase velocity estimates of both Rayleigh and Love waves at higher frequencies compared to classical processing approaches. This issue was beyond the scope of the ESG2006 noise blind test and could not be fully addressed since participants used different array layouts. We simulate in this paper seismic ambient noise for one of the models proposed within the ESG2006 noise blind test for an array configuration that is suitable for correlation techniques. As a main issue in the convergence of the correlation technique to the Green's function is the influence of the temporal and spatial distribution of noise sources, we computed noise synthetics in section 3.2.3 for both isotropic and unidirectional spatial distribution of sources. In sections 3.2.4 and 3.2.5, FK and SPAC analysis as well as noise correlation analysis are then applied on the vertical component. Results obtained by different methods are compared to the true dispersion curves and to one another. Finally, we conclude in section 3.2.6 on practical considerations for the implementation of noise correlation techniques for subsurface passive imaging.

► 3.2.3 Ambient noise simulation and array configuration

Ambient noise simulation

Ambient noise was simulated in a complex shallow structure with strong impedance contrast and complex layering including lowvelocity zones. This ground profile is similar to model N102 proposed within the framework of ESG2006 Noise Blind test [Cornou et al., 2006]. The compressional (V_p) and shear wave (V_s) profiles of this model as well as corresponding dispersion curves are displayed in Fig. 3.1.

Regarding noise synthetics generation, noise sources were approximated by subsurface forces located at 0.5-m depth with random force orientation and amplitude [Moczo & Kristek, 2002]. Distribution of sources is random in time. In this study, two distinct distributions of source locations were considered: an isotropic and a unidirectional distribution as depicted in Fig. 3.2. The source

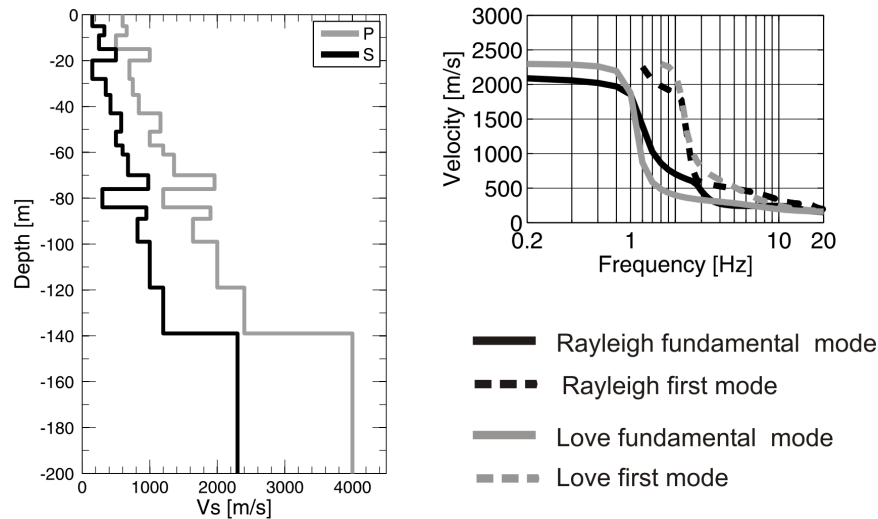


Figure 3.1: (Left-hand panel) P- and S-wave velocity profiles (grey and black curves, respectively); (right-hand panel) corresponding Rayleigh and Love waves dispersion curves for the fundamental and first higher modes.

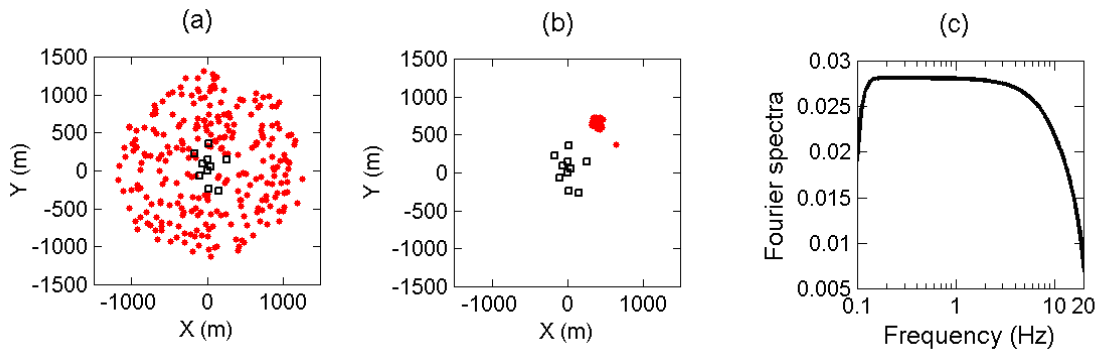


Figure 3.2: (a) Source locations (red dots) for isotropic noise. (b) Source locations (red dots) for directional noise. Receiver locations are indicated by black squares. (c) Fourier amplitude spectra of the delta-like signal used as source time function.

time function employed at each point location is a delta-like signal with a frequency Fourier amplitude spectrum flat from 0.1 to 20 Hz. Computation up to 20 Hz of the associated wave field has then been performed using the wavenumber-based technique of Hisada [1994, 1995] for 1-D horizontally layered structures. In case of isotropic source distribution, noise data set (hereafter called isotropic noise) is composed of three sets of 30 min of noise synthetics, each of them implying source shots at 74–268 different locations excited at 60 different random times, while for the directional source distribution one data set of 30 min was considered (hereafter called directional noise). Such duration of noise recordings is similar to duration usually considered in real world experiment. Fig. 3.3 displays example of noise synthetics.

Choice of an optimized seismic array layout

This paper goal is to compare correlation technique, SPAC and FK techniques. To do so, the array was designed to be suitable for correlation technique as well as for SPAC and FK techniques and

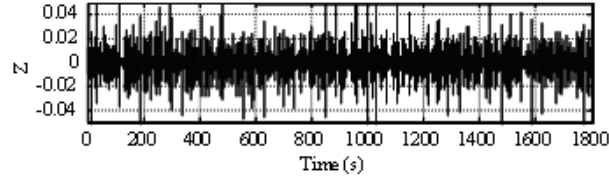


Figure 3.3: A typical signal (particle velocity) for the vertical component.

to mimic classical ambient noise studies. Following criteria were followed.

1. The number of receivers is restricted to 10.
2. As the noise is not necessarily isotropic, a large number of azimuths have to be scanned by the different pairs of receivers to average over angle distribution.
3. The spatial extension of the array has to be adapted to the targeted wavelength interval. Expecting a 500 m s^{-1} surface wave velocity at 5 Hz, the wavelength of interest is centred around 100 m.
4. The distance between neighbouring stations must sample a range interval with minimal gaps and over sampling, covering between 1 and 5 wavelengths to provide enough resolution in the F-K domain.

The spiral shape of the array shown in Fig. 3.4 fulfils the above criteria. This array roughly presents an even interreceiver distance range sampling from 72 to 637 m.

Array response

Any seismic array configuration can be considered as a discrete spatial sampling of the continuous seismic wavefield in space and time. The sampling theorem then holds and the short-wavelength part of the wavefield cannot be recovered uniquely (spatial aliasing). For linear equidistantly spaced sensor arrays, the relation between the interstation distance d_{min} of neighbouring stations and the spatial Nyquist frequency λ_{Nyq} is such that each wavelength needs to be sampled (equidistantly) by at least two discrete sampling locations:

$$\lambda_{Nyq} = \lambda_{min} = 2d_{min} \quad (3.4)$$

The resolution capability of a seismic array is defined as the ability to separate two waves propagating at closely spaced wavenumbers. This resolution is related to the maximum interstation distance d_{max} , that is, to the array aperture:

$$\lambda_{max} = d_{max} \quad (3.5)$$

Eq. (3.5) is in a strict sense only valid for 1-D array layouts and the conventional beamforming algorithm [Burg, 1964, Lacoss et al., 1969]. The sensor geometry having in general a 2-D irregular shape, both the aliasing condition and the resolution capability depend on the direction of

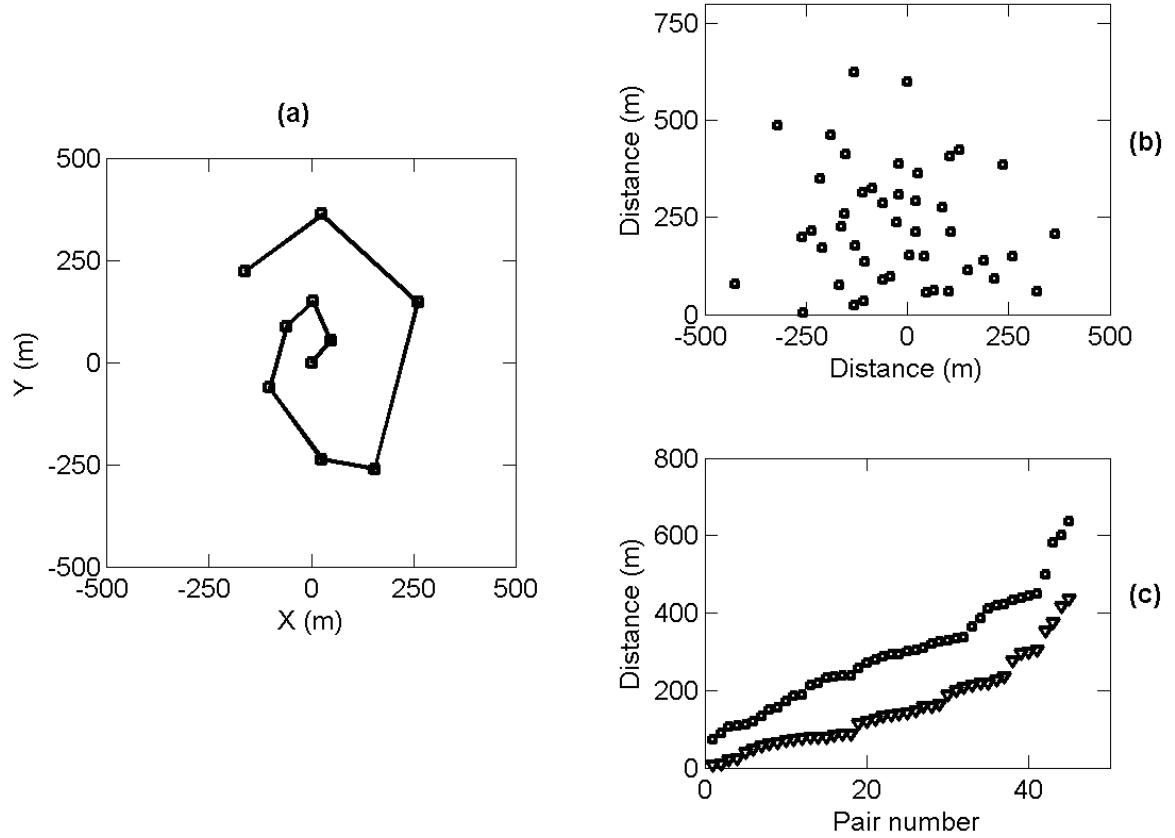


Figure 3.4: (a) array layout; (b) azimuthal distribution of receiver pairs; (c) Range sampling for receiver pairs obtained using the array in (a) (circles). Triangles correspond to the projected distribution of ranges according to a directional noise with a backazimuth of 61° from the North.

the incident wavefield and the effective smallest/largest interstation distance along the wave propagation direction [Henstridge, 1979, Asten & Henstridge, 1984, Gaffet et al., 1998, Ohrnberger, 2005, Asten, 2006, Okada, 2006, Wathelet et al., 2008]. Therefore, for arbitrary 2-D array geometries there is no simple analytic expression which relates the array shape with its spatial aliasing or resolution limits. Despite this fact, observation and numerical evaluation of the array response function have shown that the wavelength limits ($\lambda_{min}, \lambda_{max}$) as derived above can serve as a first order proxy to specify the resolution capabilities of a seismic array. However, both high-resolution frequency-wavenumber (hereafter called HRFK) techniques [Capon, 1969] as well as the SPAC technique [Aki, 1957] show improved resolution capability compared to the conventional beamformer. While for the highresolution FK method it has been repeatedly reported [e.g. Woods & Lintz, 1973, Asten & Henstridge, 1984, Tokimatsu, 1995, Satoh et al., 2001, Okada, 2003] that the longest resolvable wavelength λ_{max} is around three to six times longer than for the conventional technique ($\lambda_{max} \approx 3 - 6 d_{max}$), many authors report a superior performance for the SPAC technique compared to HRFK for the longer wavelength limit [e.g. Okada, 2003, Cornou et al., 2006]. According to many studies [Horike, 1985, Miyakoshi, 1996, Asten et al., 2004] the longest resolvable wavelength λ_{max} is in the order of 10–15 times the radius of the array configuration ($\lambda_{max} \approx 5 - 7.5 d_{max}$).

Since we use SPAC, HRFK and correlation techniques in this paper, we define for the array

resolution capability the following wavelength limits ($\lambda_{min} = 2d_{min}$, $\lambda_{max} = 3d_{max}$).

► 3.2.4 Phase velocity dispersion curves obtained from HRFK and SPAC analysis

High-resolution frequency-wavenumber (HRFK) analysis

We used the HRFK technique [Capon, 1969] as implemented in the sesarray software (<http://www.geopsy.org/>, Wathelet et al. [2008]). Operating with sliding time windows and narrow frequency bands, this method provides the wave propagation parameters (azimuth and slowness as a function of frequency) of the most coherent plane wave arrivals. The wave propagation on the vertical component was measured by using 200 frequency bands between 0.5 and 20 Hz. The central frequency of each band was selected to be equally spaced in logarithm scale. A fraction of the central frequency f_c defined the frequency bandwidth ($0.97f_c - 1.03f_c$). We selected the time window length as 300 times the central period corresponding to the analysed frequency band f_c . Fig. 3.5 displays the normalized histograms of the phase velocity estimates as a function of frequency for both noise data sets (isotropic and directional noise). Consistently with the resolution capabilities of the array, measurable phase velocities lie between resolution ($\lambda_{max} = 3d_{max}$) and aliasing limits of the array ($\lambda_{min} = 2d_{min}$). Consistently also with the known capability of FK technique that performs better for directional wave propagation, the first higher mode branch of Rayleigh waves dispersion is well retrieved for directional noise.

Spatial autocorrelation (MSPAC) analysis

Spatially average autocorrelation coefficients were computed on the vertical component following the modified vertical SPAC method (MSPAC) developed by Bettig et al. [2001] and implemented by Wathelet et al. [2004]. This method is particularly suitable here since it allows arbitrary array layouts. Autocorrelation coefficients were estimated by using sliding time windows as being three hundred times the central period corresponding to the analysed frequency band f_c . As for FK analysis we use 200 frequency bands between 0.5 and 20 Hz. Histograms of all possible phase velocities corresponding to measured autocorrelation coefficients are displayed in Fig. 3.5 for both isotropic and directional noise data sets. In case of isotropic noise, SPAC provides good estimates of phase velocities in between the array limits. For directional noise, SPAC estimates are deteriorated compared to results derived from the isotropic noise case due to the sparse distribution of interstation azimuths for such directional sources. We also observe that SPAC fails in identifying higher modes which is an inherent limitation of classical SPAC technique [e.g. Okada, 2003].

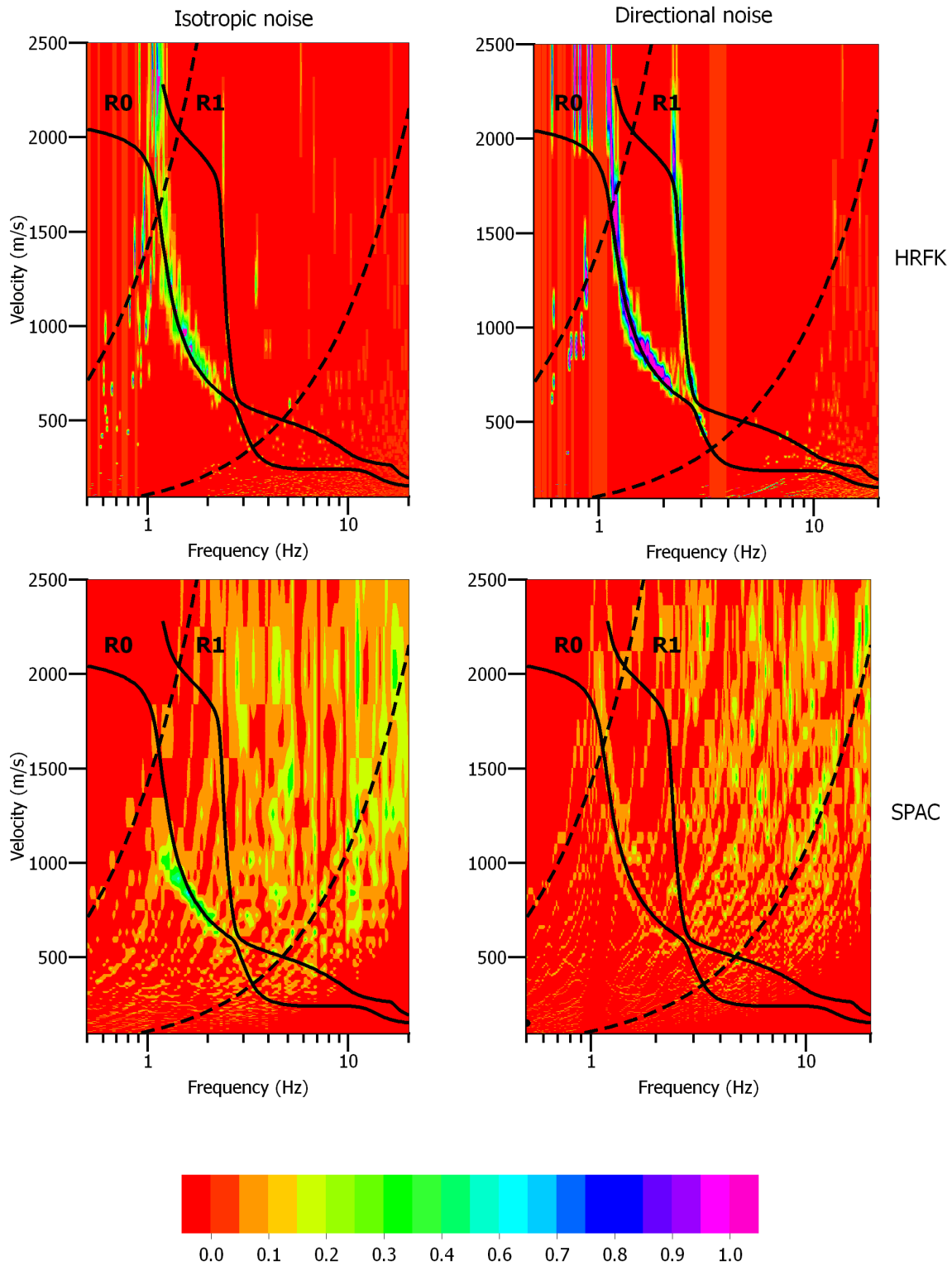


Figure 3.5: (Top) HRFK frequency-velocity normalized histograms for isotropic and directional noise; (Bottom) Normalized histograms derived from SPAC analysis for both isotropic and directional noise. Limits of array response given the following wavelength limits [$\lambda_{min} = 2d_{min}$, $\lambda_{max} = 3d_{max}$] are indicated by dashed lines. Full lines correspond to theoretical Rayleigh fundamental (R0) and first higher mode (R1).

► 3.2.5 Phase velocity dispersion curves obtained from noise correlations slantstack (NCSS)

Principle

It has been demonstrated both theoretically and experimentally that the Green's function between two receivers can be retrieved from the cross-correlation of isotropic noise records. This technique is applied here to synthetic noise to obtain the Green's function for each receiver pair of a seismic array. For a displacement field $u(\vec{r}, t)$ measured simultaneously at two receiver locations \vec{r}_1 and \vec{r}_2 , we define the normalized cross correlation over a recording time window T as:

$$C_{1,2}(\tau) = \int_0^T u(\vec{r}_1, t) u(\vec{r}_2, t + \tau) dt \Bigg/ \sqrt{\int_0^T u^2(\vec{r}_1, t) dt \int_0^T u^2(\vec{r}_2, t) dt} \quad (3.6)$$

The denominator in (3.6) is a normalization factor that practically mitigates local station effects and helps in determining the actual coherence of the noise field between the stations. Indeed, the autocorrelation $C_{1,1}(\tau)$ has a maximum of 1 while strongly decorrelated noise signals (*i.e.*, when stations are too far away or when the incoming noise is incoherent) do not exhibit any peak in the normalized noise correlation function but a random temporal signal with a variance of $1/(2T\Delta\omega) \ll 1$, when dealing with frequency-limited noise sources of bandwidth $\Delta\omega$ [Sabra et al., 2005, Weaver & Lobkis, 2005].

For each pair of the three-component receivers, the nine-component correlation tensor is constructed by using each combination of the receiver components in the cross-correlation process. Transverse and radial components are relative to the azimuthal direction of receiver pair and are computed for each pair separately by projecting the North and East components towards transverse and radial directions. In the following only the vertical-vertical (Z-Z) component of the correlation tensor will be studied for a fair comparison with SPAC and HRFK methods. Rayleigh waves are expected to be mainly observed on Z-Z and R-R components. Transverse-transverse (T-T) components should also provide useful information on Love waves, as mentioned later on. Other components of the correlation tensor do not provide direct information to measure dispersion curves.

As the medium is horizontally stratified, the extracted Green's function only depends on the distance between receivers used for computing correlation. Correlation functions obtained for all receiver pairs are then plotted versus distance to construct a seismic section, as if a classical active seismic experiment was carried out with a linear array. Phase velocity dispersion curves are subsequently obtained by applying classical FK transform to the reconstructed seismic section. For sake of conciseness we will refer hereafter this above-mentioned algorithm as noise correlation slantstack technique (NCSS).

For the analysis of both isotropic and directional noise data sets that will be presented in the following sections, synthetics are preprocessed to whiten their frequency spectrum between 0.5 and 20 Hz in order to limit the impact of high-frequencies damping in the medium. Amplitude of synthetics is also made one bit (*i.e.*, only the sign of the time-domain whitened signal is kept, see Larose et al. [2004]) to get rid off temporal variation of noise sources amplitude.

Analysis of isotropic noise

A seismic section obtained for each receiver pair from the crosscorrelation of the Z-component noise synthetics is plotted in Fig. 3.6(a). Contrary to classical seismic sections, the obtained seismic section does not exhibit equally spaced Green's functions in spatial domain, the minimum distance between two Green's functions being $d_{min} = 0.3$ m and the maximum distance being $d_{max} = 637.4$ m. Fig. 3.7(a) displays the frequency-velocity diagrams derived by applying classical FK analysis to the Z-Z seismic section. Limits of the array response are computed according to minimum and maximum distances (d_{min}, d_{max}) previously defined. Rayleigh waves' dispersion is nicely retrieved on this diagram.

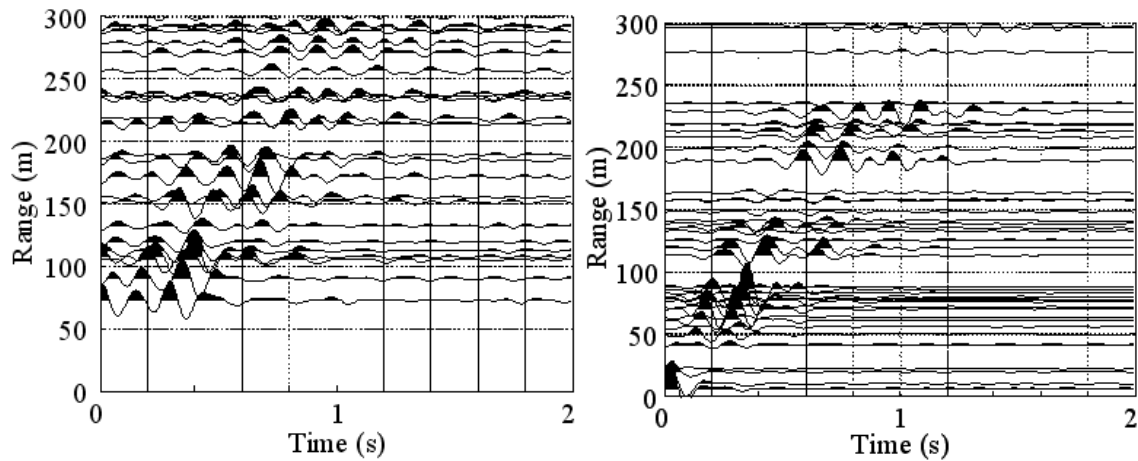


Figure 3.6: (a) Zoom between 0 and 300 m of the seismic section obtained for the Z-Z component of the correlation tensor accumulating all receiver pairs. (b) The same section obtained in the case of directional noise. The signal-to-noise ratio is higher despite shorter records in time, as all noise sources coherently contribute to the Green's function reconstruction.

Analysis of directional noise

In the case of directional noise, the correlation tensor is no longer connected to the Green's tensor since the wave arrival between each receiver pair depends on the noise directivity angle. However, assuming a 1-D medium, the correlation for each receiver pair is the Green's function along the direction of incident wave. If the direction of incident wave is known, it is then possible to account for such apparent Green's function by defining, for each receiver pair, an effective distance as being the projection of distance between receivers along the noise direction (or the difference in source-receiver distances if the source is in the near field). The noise correlation then converges to the Green's function related to this effective distance. The projection along the noise direction modifies each receiver pair distance, and thereafter resolution capabilities of the array. In this paper, direction of noise is first determined by applying a matched field processing [Baggeroer et al., 1993]. The average source backazimuth is found to be 61° . (Fig. 3.8) and effective distances are computed together with Green's function. As illustrated in Fig. 3.4(c), the minimum and maximum distances between receiver pair are reduced to 6 and 436 m, respectively. The seismic section obtained from the Z-Z correlation tensor is shown in Fig. 3.6(b). Compared to the isotropic noise data set, the signal-to-noise ratio is strongly increased as each source is now coherently contributing to the

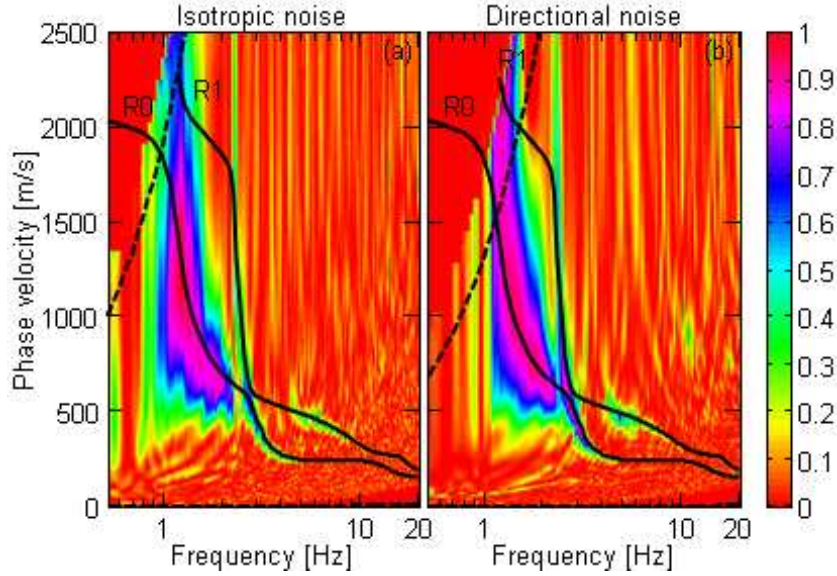
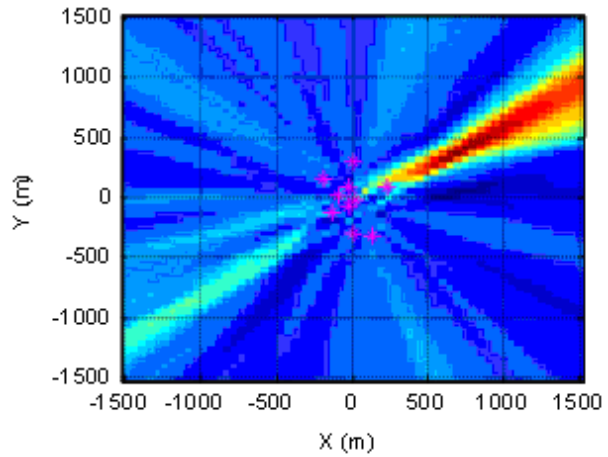


Figure 3.7: Frequency-velocity diagrams for NCSS applied to the Z–Z component of the correlation tensor in the case of isotropic (a) and directional (b) noise. The two first theoretical Rayleigh modes (R0 and R1 modes) are indicated by full lines. Limits of array response given the following wavelength limits [$\lambda_{min} = 2d_{min}$, $\lambda_{max} = 3d_{max}$] are indicated by dashed lines (low velocity limits corresponding to λ_{min} is very close to the frequency axis).

Green's function reconstruction when distances are projected along the noise direction. Frequency-velocity diagram derived from the FK processing is displayed in Fig. 3.7(b). Since the signal-to-noise ratio of seismic sections has been increased, dispersion curve of Rayleigh waves are better retrieved than for the isotropic noise data set.

Figure 3.8: Matched field processing indicates impinging sources (red-colour region) propagating with a backazimuth of N61°. (Y positive direction indicates North and X positive direction indicates East). Purple stars indicate station locations.



► 3.2.6 Results comparison

Fig. 3.9 displays phase velocity estimates as a function of frequency obtained for both isotropic and directive noise data sets and derived by HRFK and NCSS. In order to better quantify the ability of each technique in estimating phase velocities, we have displayed in Fig. 3.10 the relative deviation of measured phase velocities from the true velocities (see Cornou et al. [2006] for details about

Table 3.1: Minimum (d_{min}) and maximum (d_{max}) distances serving as array-resolution proxies.

Sources	d_{min}		d_{max}	
	SPAC/HRFK	NCSS	SPAC/HRFK	NCSS
Random (m)	72.3	0.3	637.4	637.4
Directional (m)	72.3	0.18	637.4	436.2

For SPAC and HRFK techniques, d_{min} stands for the minimum distance between neighbourhood stations and d_{max} stands for the array aperture. For noise correlation slantstack (NCSS) technique, d_{min} and d_{max} stands for the minimum and maximum distances between Green's functions observed on the reconstructed seismic profile.

the plotting procedure). The most striking feature is the ability of NCSS technique in estimating Rayleigh wave phase velocities at higher frequency than HRFK technique for both directional and isotropic noise. Indeed, HRFK provides phase velocity estimates between 1 and 3 Hz, while NCSS technique provides estimates between 1 and 12 Hz. For the array layout used in this paper, the superiority of the NCSS technique mainly lies in the reconstruction of a 1-D seismic section with spatial sampling between reconstructed Green's functions smaller than the minimum distance between neighbourhood receivers. As a consequence, the spatial Nyquist frequency is reduced and shorter wavelength parts of the wavefield can thus be recovered compared to classical array approaches. On the opposite, the previously defined array proxy-capability to resolve the longest wavelength parts of the wavefield is similar for both approaches: the maximum distance between receivers for both HRFK and SPAC analysis or between Green's functions for the reconstructed seismic profile are indeed similar, except when specific processing is performed to account for directive sources. In such a case, projection of distances between receivers along the direction of the incident wave field leads to effective distances smaller than the initial distance sampling, which deteriorates array capability for longest wavelengths.

Table 3.1 summarizes the minimum and maximum distances for both HFRK/SPAC and NCSS approaches that control the array proxy-capabilities as defined in this study. One has to mention however that, besides considerations based on the theoretical array proxycapabilities, the actual performance of techniques to resolve lowest frequencies (*i.e.*, below 1 Hz) can not be fully addressed here due to the high pass filtering effect of the sediment layer leading to the vanishing of vertical spectral energy below 1 Hz [see Cornou et al., 2006].

Regarding effects of noise source spatial distribution, phase velocities of Rayleigh first mode (R1) are better retrieved for both techniques when noise is directional (Figs 3.9 and 3.10). This is due to the better ability of FK analysis for a directive source and to the improved signal-to-noise ratio for NCSS technique as mentioned in the previous section. For HRFK analysis and both directional and isotropic noise, phase velocities estimates are very close to the theoretical dispersion curves, the median value of absolute relative deviation ranging between 3 and 8 per cent. The estimates are however slightly biased to higher velocities, especially at longest wavelengths (below 1.2 Hz), as a consequence of insufficient array resolution [see Cornou et al., 2006]. For correlation technique and directional noise, fundamental Rayleigh wave phase velocities are particularly well retrieved above 1.2 Hz: the median value of absolute relative deviation is less than 1 per cent for

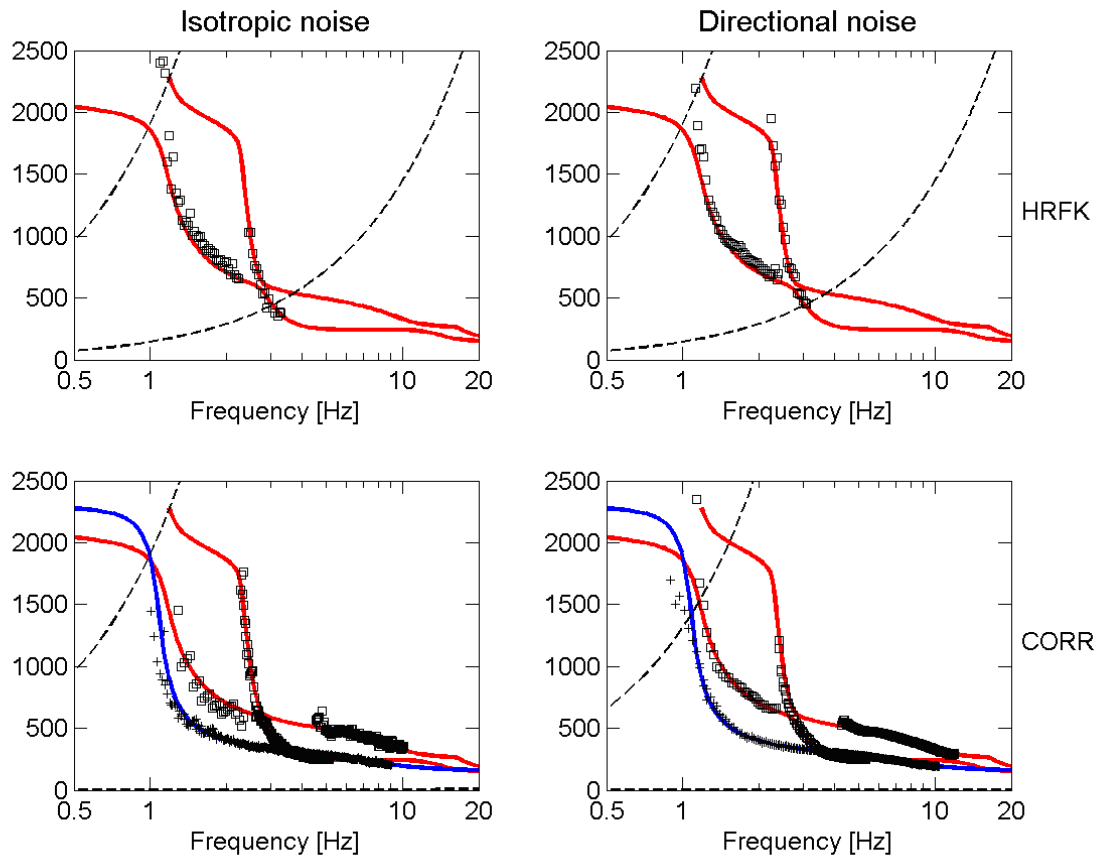


Figure 3.9: (Top) Phase velocity estimates of fundamental and higher modes of Rayleigh waves (black squares) derived from frequency-velocity normalized histograms (Fig. 3.5) using HRFK analysis. (Bottom) Phase velocity estimates of fundamental and higher modes of Rayleigh (black squares) and Love waves (black crosses) derived from frequency-velocity diagrams (Fig. 3.7) using NCSS. Limits of array response given the following wavelength limits [$\lambda_{min} = 2d_{min}$, $\lambda_{max} = 3d_{max}$] are indicated by dashed lines. Theoretical dispersion curves of fundamental and higher modes of Rayleigh and Love waves are indicated by red and blue lines, respectively.

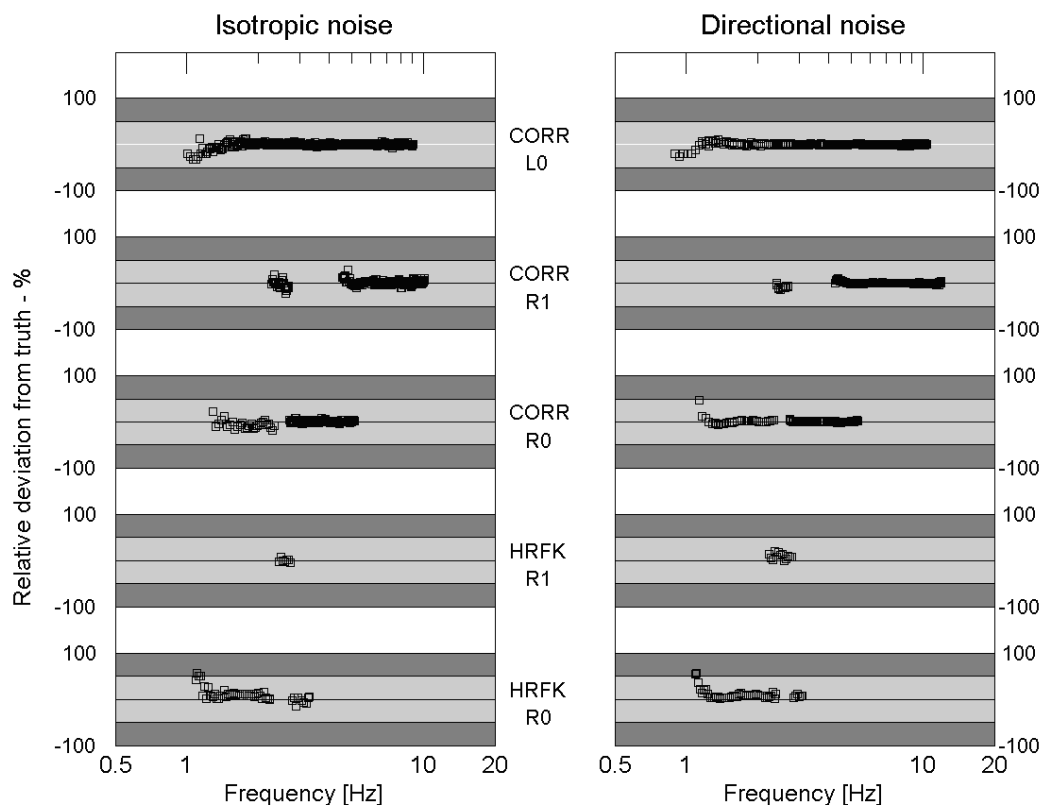


Figure 3.10: Relative deviation of phase estimates (squares) from the true dispersion curves obtained for both isotropic (right-hand panel) and directional (left-hand panel) noise data sets. Methods used to measure phase velocities are indicated (CORR stands for correlation technique, while HRFK stands for High-Resolution frequency-wavenumber analysis) together with the surface wave modes (R0 and R1 stands for fundamental and first higher Rayleigh modes, respectively; L0 stands for fundamental Love mode). The relative deviation is given in percent; grey bars correspond to the following deviation range: [-50 50 per cent], [-100 -50 per cent], [50 100 per cent].

Rayleigh fundamental and first modes), while deviation to higher velocities can be observed below 1.2 Hz. For isotropic noise, phase velocity estimates are more scattered, especially for the fundamental Rayleigh mode below 2.7 Hz. However, median value of absolute relative velocity deviation is 2 per cent above 2.7 Hz and 7.6 per cent below 2.7 Hz. Such scattering can be explained by the low signal-to-noise ratio observed on the reconstructed seismic profile compared to the one derived from directional noise (Fig. 3.6). Considering longer time series would certainly improve the signal-to-noise ratio and ensure that correlation converges towards more robust Green's functions.

Phase velocity dispersion curves may also be extracted from R–R and T–T components of the cross-correlation tensor. This is a clear advantage of the NCSS technique with respect to HRFK. Rayleigh waves are expected on the R–R component and Love waves on the T–T component. Using R–R component is useful to improve estimate of Rayleigh dispersion curve by providing an independent measure. T–T component of the correlation tensor offers a way to measure Love wave dispersion curves with a high accuracy as they are separated from Rayleigh waves. Figs 3.9 and 3.11 show Love wave dispersion curves obtained using NCSS analysis for both isotropic and directive noise source data sets. Phase velocity estimates are very close to the theoretical dispersion curves outlining the ability of such technique to also retrieve information on Love waves propagation.

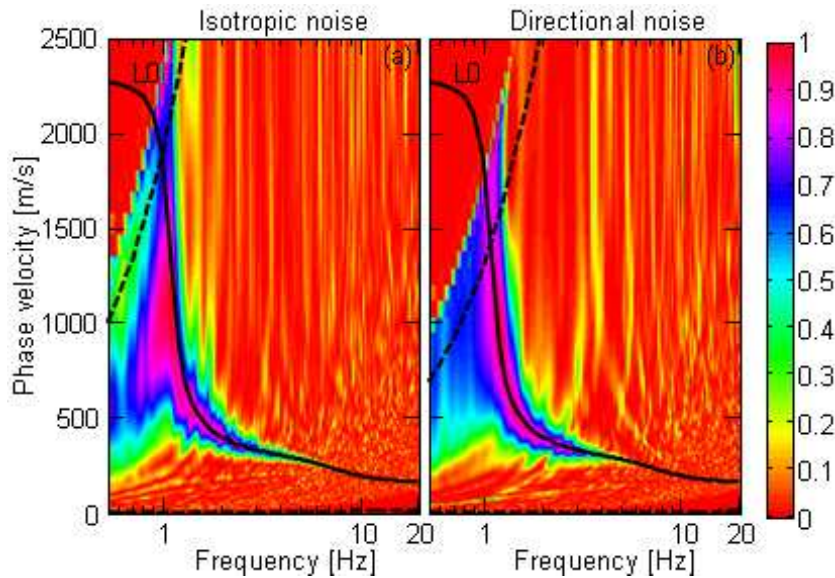


Figure 3.11: Frequency-velocity diagrams for the T–T component of the correlation tensor in the case of isotropic (a) and directional (b) noise. The fundamental theoretical mode of Love waves (L0) is indicated in full line. Limits of array response given the following wavelength limits [$\lambda_{min} = 2d_{min}$, $\lambda_{max} = 3d_{max}$] are indicated by dashed lines (low velocity limits corresponding to λ_{min} is very close to the frequency axis).

► 3.2.7 Conclusion

In this paper, we have simulated ambient seismic noise for both isotropic and directional spatial distribution of noise sources on a given array layout. Noise synthetics were then analysed using HRFK, SPAC technique and NCSS technique. Results outlined that NCSS provides phase velocity estimates of Rayleigh waves within a wider frequency band (between 1 and 12 Hz) than classical

techniques (between 1 and 3 Hz). We explain this superiority of NCSS technique by the processing step that consists in deriving a 1-D seismic section composed of reconstructed Green's functions with an effective spatial sampling smaller than the absolute minimum interstation distance. As a result, the effective spatial Nyquist frequency is smaller for correlation techniques than for classical approaches. Applying a frequency-wavenumber transform to the reconstructed seismic section allows the NCSS technique to recover shorter wavelengths of the wavefield than classical array approaches.

NCSS technique is a very promising tool for subsurface structure imaging as the estimates only present here 1–2 per cent error, compared to 3–8 per cent for HRFK. Furthermore, Love wave dispersion curves can also be accurately obtained with no extra effort. Here, the choice of the array geometry is obviously a key factor for retrieving dispersion curves at higher frequencies than classical techniques. The use of typical array geometries (circles, triangles, L-shape arrays) commonly used in ambient noise array technique would certainly not lead to such good results at high frequency since the spatial sampling of such arrays do not span a wide range of distances. Further studies should be performed in order to derive the typical most appropriate array layouts suitable for NCSS techniques.

In this paper, 30 min of noise recording is enough to measure accurately Rayleigh dispersion curves from NCSS technique in the targeted frequency band. Convergence and sensitivity studies have however to be carried out to adjust the recording time to practical cases. Besides, NCSS technique performs better in the case of directional noise since it takes advantage of noise directivity to increase signal-to-noise ratio. In case of several non-isotropic noise directions, NCSS technique is no more suitable as several possible projected ranges exist for one receiver pair. The cross-correlation is then a mix between different Green's functions.

Finally, we used a here FK method, as one among many processing techniques used in classical active surveys, to measure dispersion curves from the reconstructed seismic profile. The possibility to retrieve other information than dispersion curves (e.g. ellipticity, attenuation, etc.) from cross-correlation has to be explored.

► Acknowledgements

Noise synthetics were performed at the Service Commun de Calcul Intensif (SCCI) of Grenoble, France observatory (OSUG). We thank the two anonymous reviewers for their constructive comments.

Bibliographie chapitre 3

Aki, K., 1957. Space and time spectra of stationary stochastic waves with special reference to microtremors, *Bulletin of the Earthquake Research Institute*, **35**, 415–456.

Aki, K., 1965. A note on the use of microseisms in determining the shallow structures of the Earth's crust, *Geophysics*, **30**(4), 665–666.

Asten, M. W., 2006. On bias and noise in passive seismic data from finite circular array data processing using SPAC methods, *Geophysics*, **71**(6), V153–V162.

Asten, M. W. & Henstridge, J. D., 1984. Array estimators and use of microseisms for reconnaissance of sedimentary basins, *Geophysics*, **49**, 1828–1837.

Asten, M. W., Dhu, T., & Lam, N., 2004. Optimised array design for microtremor array studies applied to site classification; observations, results and future use, in *Proceedings of the 13th World Conference of Earthquake Engineering, Vancouver*.

Baggeroer, A. B., Kuperman, W. A., & Mikhalevsky, P. N., 1993. An overview of matched fields methods in ocean acoustics, *IEEE Journal of Ocean Engineering*, **18**(4), 401–424.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophysical Journal International*, **169**(3), 1239–1260.

Bettig, B., Bard, P.-Y., Scherbaum, F., Riepl, J., Cotton, F., Cornou, C., & Hatzfeld, D., 2001. Analysis of dense array noise measurements using the modified spatial auto-correlation method MSPAC – application to the Grenoble area, *Bollettino di Geofisica Teorica e Applicata*, **42**(3-4), 281–304.

Burg, J. P., 1964. Three-dimensional filtering with array of seismometers, *Geophysics*, **29**, 693–713.

Campillo, M., 2006. Phase and correlation in ‘random’ seismic fields and the reconstruction of the Green function, *Pure and Applied Geophysics*, **163**(2-3), 475–502.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**, 547–549.

Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proceedings of the IEEE*, **57**(8), 1408–1418.

Chávez-García, F. J., Rodríguez, M., & Stephenson, W. R., 2005. An alternative approach to the SPAC analysis of microtremors: exploiting stationarity of noise, *Bulletin of the Seismological Society of America*, **95**(1), 277–293.

Cho, I., Tada, T., & Shinozaki, Y., 2004. A new method to determine phase velocities of Rayleigh waves from microseisms, *Geophysics*, **69**(6), 1535–1551.

- Cho, I., Tada, T., & Shinozaki, Y., 2006. A generic formulation for microtremor exploration methods using three-components records from a circular array, *Geophysical Journal International*, **165**, 236–258.
- Cornou, C., Ohrnberger, M., Boore, D. M., Kudo, K., & Bard, P.-Y., 2006. Using ambient noise array techniques for site characterizations: results from an international benchmark, in *Proceedings of the 3rd International Symposium on the Effects of Surface Geology on Seismic Motion, Grenoble, 29 August – 01 September 2006*, LCPC Editions, paper NBT.
- Derode, A., Larose, E., Campillo, M., & Fink, M., 2003. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Applied Physics Letters*, **83**(15), 3054–3056.
- Duvall, T. L., Jefferies, S. M., Harvey, J. W., & Pomerantz, M. A., 1993. Time distance helioseismology, *Nature*, **362**, 430–432.
- Gaffet, S., Larroque, C., Deschamps, A., & Tressols, F., 1998. A dense array experiment for the observation of waveform perturbations, *Soil Dynamics and Earthquake Engineering*, **17**, 475–484.
- Gilles, P. M., Duvall, T. L., Scherrer, P. H., & Bogart, R. S., 1997. A subsurface flow of material from the Sun's equator to its poles, *Nature*, **390**, 52–54.
- Gouédard, P., Roux, P., & Campillo, M., 2006. Dispersion curves measurement using noise cross-correlation technique, in *Proceedings of the 3rd International Symposium on the Effects of Surface Geology on Seismic Motion, Grenoble, 29 August – 01 September 2006*, LCPC Editions, paper N06.
- Gouédard, P., Cornou, C., & Roux, P., 2008. Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique, *Geophysical Journal International*, **172**(3), 971–981.
- Henstridge, J. D., 1979. A signal processing method for circular arrays, *Geophysics*, **44**, 179–184.
- Hisada, Y., 1994. An efficient method for computing Green's functions for a layered halfspace with sources and receivers at close depths, *Bulletin of the Seismological Society of America*, **84**, 1456–1472.
- Hisada, Y., 1995. An efficient method for computing Green's functions for a layered halfspace with sources and receivers at close depths (part 2), *Bulletin of the Seismological Society of America*, **85**(4), 1080–1093.
- Horike, M., 1985. Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas, *Journal of Physics of the Earth*, **33**(2), 59–96.
- Kanai, K. & Tanaka, T., 1954. Measurement of microtremor, *Bulletin of the Earthquake Research Institute*, **32**, 199–209.
- Lacoss, R. T., Kelly, E. J., & Toksöz, M. N., 1969. Estimation of seismic noise structure using arrays, *Geophysics*, **34**(1), 21–38.

- Larose, E., Derode, A., Campillo, M., & Fink, M., 2004. Imaging from one-bit correlation of wide-band diffuse wavefield, *Journal of Applied Physics*, **95**(12), 8393–8399.
- Larose, E., Lobkis, O. I., & Weaver, R. L., 2006a. Passive correlation imaging of a buried scatterer, *The Journal of the Acoustical Society of America*, **119**(6), 3549–3552.
- Larose, E., Margerin, L., Derode, A., van Tiggelen, B. A., Campillo, M., Shapiro, N. M., Paul, A., Stehly, L., & Tanter, M., 2006b. Correlation of random wavefields: an interdisciplinary review, *Geophysics*, **71**(4), SI11–SI21.
- Lin, F., Ritzwoller, M. H., Townend, J., Bannister, S., & Savage, M., 2007. Ambient noise Rayleigh wave tomography of New Zealand, *Geophysical Journal International*, **170**(2), 649–666.
- Louie, J., 2001. Faster, better: shear-wave velocity to 100 meters depth from refraction microtremor array, *Bulletin of the Seismological Society of America*, **91**, 347–364.
- Miyakoshi, K., 1996. A range of wavelengths possible to estimate phase velocities of surface waves in microtremors, in *Proceedings of the 94th SEGJ Conference, Society of Exploration Geophysicists, Japan*, pp. 178–182, (in Japanese).
- Moczo, P. & Kristek, J., 2002. Fd code to generate noise synthetics, in *Sesame report*, no. D02.09, available on <http://sesame-fp5.obs.ujf-grenoble.fr/>.
- Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface using microtremors on the ground surface, *Quarterly Reports of Railway Technological Institute of Tokyo*, **30**, 25–33.
- Nogoshi, M. & Igarashi, T., 1971. On the amplitude characteristics of microtremor (part 2), *Journal of the Seismological Society of Japan*, **24**, 26–40, (in Japanese with English abstract).
- Ohori, M., Nobata, A., & Wakamatsu, K., 2002. A comparison of ESAC and FK methods of estimating phase velocity using arbitrary shaped microtremor array, *Bulletin of the Seismological Society of America*, **92**(6), 2323–2332.
- Ohrnberger, M., 2005. Report on the FK/SPAC capabilities and limitations, in *SESAME Deliverable*, no. D19.06, pp. 43, <http://sesame-fp5.obs.ujf-grenoble.fr/Delivrables/Del-D19-Wp06.pdf>.
- Okada, H., 2003. The microseismic survey method, in *Geophysical Monographs*, vol. 12, Society of Exploration Geophysicists of Japan, translated by Koya Suto, Society of Exploration Geophysicists, Tulsa.
- Okada, H., 2006. Theory of efficient array observations of microtremors with special reference to the SPAC method, *Exploration Geophysics*, **37**, 73–85.
- Pedersen, H. A., Krüger, F., & the SVEKALAPKO Seismic Tomography Working Group, 2007. Influence of the seismic noise characteristics on noise correlations in the Baltic Shield, *Geophysical Journal International*, **168**(1), 197–210.

- Roux, P. & Kuperman, W. A., 2004. Extracting coherent wavefronts from acoustic ambient noise in the ocean, *The Journal of the Acoustical Society of America*, **116**(4), 1995–2003.
- Sabra, K. G., Gerstoft, P., Roux, P., & Kuperman, W. A., 2005. Surface wave tomography from microseisms in southern California, *Geophysical Research Letters*, **32**, L14311.
- Sánchez-Sesma, F. J. & Campillo, M., 2006. Retrieval of the Green function from cross correlation: the canonical elastic problem, *Bulletin of the Seismological Society of America*, **96**(3), 1182–1191.
- Sánchez-Sesma, F. J., Pérez-Ruiz, J. A., Campillo, M., & Luzón, F., 2006. Elastodynamic 2-D Green function retrieval from cross-correlation: canonical inclusion problem, *Geophysical Research Letters*, **33**, L13305.
- Satoh, T., Kawase, H., Iwata, T., Higashi, S., Sato, T., Irikura, K., & Huang, H.-C., 2001. A-wave velocity structure of the Taichung basin, Taiwan, estimated from array and single-station records of microtremors, *Bulletin of the Seismological Society of America*, **91**(5), 1267–1282.
- Shapiro, N. M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophysical Research Letters*, **31**, L07614.
- Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.
- Stehly, L., Campillo, M., & Shapiro, N. M., 2007. Travel time measurements from noise correlation: stability and detection of instrumental time-shifts, *Geophysical Journal International*, **171**(1), 223–230.
- Tokimatsu, K., 1995. Geotechnical site characterization using surface waves, in *Proceedings of the 1st International Conference of Earthquake Geotechnical Engineering*, pp. 1333–1368, Balkema.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's Function of an arbitrary inhomogeneous medium by cross-correlation, *Physical Review Letters*, **93**(25), 254301.
- Wathelet, M., Jongmans, D., & Ohrnberger, M., 2004. Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surface Geophysics*, **2**, 211–221.
- Wathelet, M., Jongmans, D., Ohrnberger, M., & Bonnefoy-Claudet, S., 2008. Array performances for ambient vibrations on a shallow structure and consequences over V_s inversion, *Journal of Seismology*, **12**(1), 1–19.
- Weaver, R. L. & Lobkis, O. I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Physical Review Letters*, **87**(13), 134301.
- Weaver, R. L. & Lobkis, O. I., 2005. The mean and variance of diffuse field correlations in finite bodies, *The Journal of the Acoustical Society of America*, **118**(6), 3447–3456.
- Woods, J. W. & Lintz, P. L., 1973. Plane waves at small arrays, *Geophysics*, **38**, 1023–1041.

Yao, H., van der Hilst, R. D., & de Hoop, M. V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps, *Geophysical Journal International*, **166**(2), 732–744.

Utilisation de sources non synchronisées

Sommaire

4.1	Small scale geophysics inversion using surface waves extracted from noise cross-correlation	106
4.1.1	Abstract	106
4.1.2	Introduction	106
4.1.3	Retrieving a seismic section from user-defined noise cross correlation	107
4.1.4	Geophysics inversion	111
4.1.5	Conclusion	112
4.2	Étude d'un milieu latéralement inhomogène	113
4.2.1	Étude des profils linéaires	114
4.2.2	Utilisation du réseau en spirale	120
4.2.3	Conclusion sur cette étude	124
4.3	Étude de la Solfatare, Italie	124
4.3.1	Présentation de la Solfatare	124
4.3.2	Étude des données actives	126
4.4	Conclusion sur l'utilisation de sources non synchronisées	129
	Bibliographie chapitre 4	131

Nous avons vu dans le chapitre précédent que les corrélations de bruit pouvaient permettre une mesure précise des courbes de dispersion pour les ondes de surface. Afin d'obtenir ces résultats, nous avons dû prendre en compte la distribution des sources de bruit dans le traitement des sismogrammes. Dans ce chapitre, l'approche utilisée est différente. Ne connaissant pas la distribution des sources de bruit ambiant, nous allons générer notre propre « bruit » d'énergie supérieure à celle du bruit ambiant. La distribution des sources peut ainsi facilement être contrôlée. Ce bruit est généré par ce que j'appellerai des « sources non synchronisées », par opposition aux sources habituellement utilisées en géophysique de petites échelles, qui doivent être parfaitement synchronisés afin de connaître le temps de référence pour la mesure des durées de propagation. Cette synchronisation des sources doit se faire avec une grande précision afin de ne pas biaiser les mesures, ce qui

complique le protocole d'acquisition des données.

Les sources non synchronisées utilisées pour cette thèse sont des pas humains. Le bruit est ainsi généré par plusieurs personnes marchant dans une zone d'extension plus ou moins importante suivant la géométrie du réseau de capteurs utilisé. D'autres sources pourraient être utilisées, suivant la gamme de fréquences utiles visées et/ou la distance de propagation voulue, les signaux de pas n'étant pas assez énergétiques pour être enregistrés pour de grandes distances de propagation.

4.1 Small scale geophysics inversion using surface waves extracted from noise cross-correlation

Cette section est une retranscription de l'article [Gouédard, Roux, & Campillo, 2008] publié au cours de ma thèse dans *The Journal of the Acoustical Society of America – Express Letters*. Cet article présente une expérience à petite échelle (profil linéaire de 14 m de long) dans laquelle nous avons appliqué une technique d'inversion à des données reconstruites par corrélation. L'originalité de cette expérience est que les sources de « bruit » utilisées sont des pas humains. L'utilisation de sources artificielles non contrôlées permet de choisir la distribution de bruit, dont on a vu qu'elle a une importance capitale dans le chapitre 2, tout en s'affranchissant des problèmes techniques de synchronisation des sources. De plus, le procédé exposé dans cet article permet des études locales en très peu de temps, puisque le déploiement, l'acquisition et l'inversion peuvent être effectués en moins d'une heure.

► 4.1.1 Abstract

Green's functions can be retrieved between receivers from the correlation of ambient seismic noise or with an appropriate set of randomly distributed sources. This principle is demonstrated in small-scale geophysics using noise sources generated by human steps during a 10-min walk in the alignment of a 14-m-long accelerometer line array. The time-domain correlation of the records yields two surface wave modes extracted from the Green's function between each pair of accelerometers. A frequency–wavenumber Fourier analysis yields each mode contribution and their dispersion curve. These dispersion curves are then inverted to provide the one-dimensional shear velocity of the near surface.

► 4.1.2 Introduction

It has been demonstrated both theoretically and experimentally that records of seismic noise, equidistributed in angle, yield the Green's functions between receiver pairs after a time-averaged cross-correlation process. Since 2002, several papers have been published on this noise cross-correlation process in geophysics applied from a hundred of meter scale to a hundred of kilometers scale [Campillo & Paul, 2003, Wapenaar, 2003, Snieder, 2004, Roux et al., 2005].

One necessary ingredient to extract Green's function from noise correlation is the angle isotropy of seismic noise or, at least, the knowledge of the noise directivity on the seismic network. This

strong hypothesis is rarely verified in practical situations when dealing with high frequency seismic noise (above 5 Hz) and small scales (less than 1 km) where noise sources have human origin. A way to get around the noise source distribution ambiguity is to perform the correlation process from a set of user-defined noise sources. When the seismic network is defined as a line of geophones, one way to deal with user-defined noise sources is to spread them in range on 10–50 wavelengths along the receiver alignment.

This technique presents several advantages. First we do not need an active device as a source. The strength of the source is often a problem when working in environments (such as an urban area, for example) where the use of potentially destructive devices such as explosives is prohibited. Here it is the duration of the record that governs the available energy of the signal. Furthermore, no synchronization of source and acquisition is required. However, using a source to perform velocity measurement between two stations is well known and currently used in seismology but it generally leads to apparent velocities which must be corrected according to the source location. Here by using a spatial distribution of sources, whose contributions are eventually averaged, we extract a response between the stations that is independent of the paths source stations, as far as the illumination is sufficient for the Green's function emergence. This condition requires in practice that the source, *i.e.*, the steps, cover a region corresponding to the “end fire lobes” of the receiver line [Roux & Kuperman, 2004]. In this case, no angular correction relying on a speculative velocity model is needed. In conclusion, extracting the Green's function from user-defined noise sources makes the correlation technique a trade-off between an active and a passive method—active since we do not rely on the natural ambient noise in the medium, but passive since the correlation process naturally achieves synchronization and averaging.

The goal of this work is to perform an experimental demonstration of the correlation process of user-defined noise sources at the meter scale using a linear array of geophones. This letter is divided into two parts. In Sec. 4.1.3, we present the experimental setup, the recording of user-defined noise sources, and the averaging correlation process that leads to the final seismic section. In Sec. 4.1.4, a classical geophysical inversion is performed from the phase velocity dispersion curves of two Rayleigh modes extracted from the seismic section. Finally, the shear velocity profile is inverted over the first 5 m.

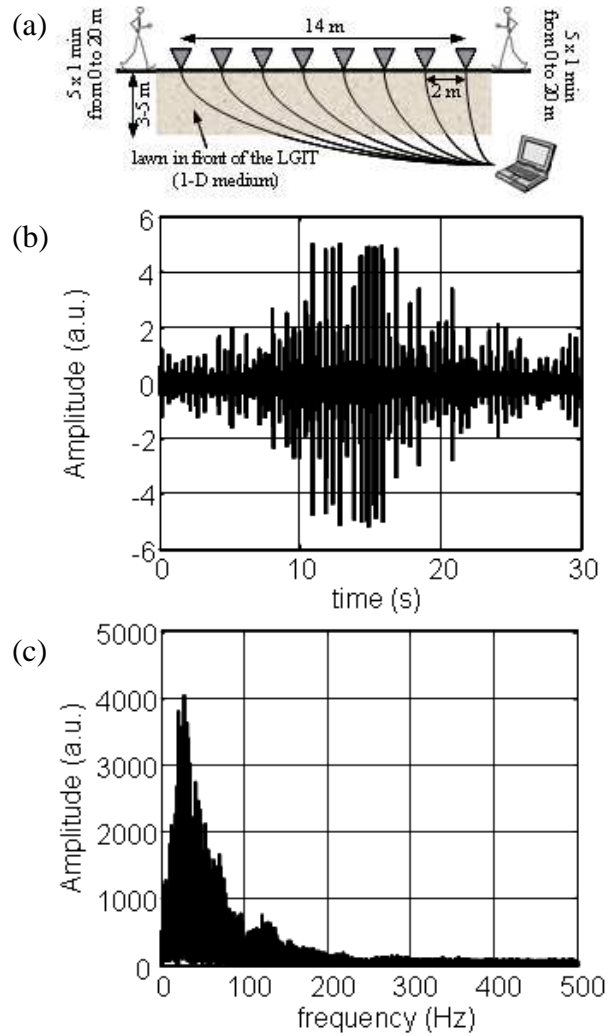
► 4.1.3 Retrieving a seismic section from user-defined noise cross correlation

A 14-m-long line of eight evenly spaced vertical accelerometers has been used to record human steps [Fig. 4.1(a)]. Two persons walked in the alignment of the accelerometers line, five times 1 min on each side, from 0 to about 30 m away from the geophone array. For a matter of convenience, the studied medium was the lawn in front of the laboratory.

The experimental setup has been designed to be versatile. The array configuration may include up to 16 one-component seismic stations. These seismic sensors could be accelerometers or geophones depending on the expected frequency bandwidth. Similarly, the array length is adjustable to the surface wave wavelength.

The main advantage of this system is to be easy and fast to setup. Our ambition was to perform

Figure 4.1: (a) Experimental setup. The acquisition system is performed from a National Instrument board that gives access to the recordings in real time. The array deployment involved eight geophones but can be adapted to various configurations using up to 16 accelerometers or geophones. (b) Typical 30-s-long time-domain recording and (c) its frequency spectrum. The foot step amplitude depends on the range to the geophone. Frequency spectrum ranges from 4 to about 150 Hz.



a complete deployment, acquisition, and inversion in approximately 30 min. The system design makes it very convenient for local and near surface measurements. Figure 4.1(b) shows a typical time-domain recorded signal with its frequency spectrum. Each event in the signal is a step whose amplitude depends on the range to the receiver. The energy spectrum of the 30-s-long record is spread up to 150 Hz. Given the frequency response of the accelerometers and the spatial extension of the array, a frequency interval ranging from 10 to 100 Hz was selected for the analysis.

Since the frequency spectrum of the steps [Fig. 4.1(c)] is not flat in the frequency interval of interest, and as correlating is mathematically equivalent to a spectrum product, only the most energizing frequencies will emerge in the correlation signal. To enlarge the effective frequency bandwidth, the spectrum of the records will be equalized in the selected frequency interval [10–100 Hz] before the correlation process. To check the robustness of the correlation process, 5 1-min long records were correlated separately for each receiver pair. Figure 4.2(a) shows, for a given pair, that the five time-domain correlations superimpose in phase. Thus correlation is robust and does not depend on the person who walked. As those correlations signals superimpose, they are stacked to increase the signal-to-noise Ratio (SNR). The residual signal observed at large lag times in the correlation (for time ~ -0.5 s) corresponds to the correlation between two consecutive steps. This signal is incoherent over the five realizations and vanishes in the 5-min average.

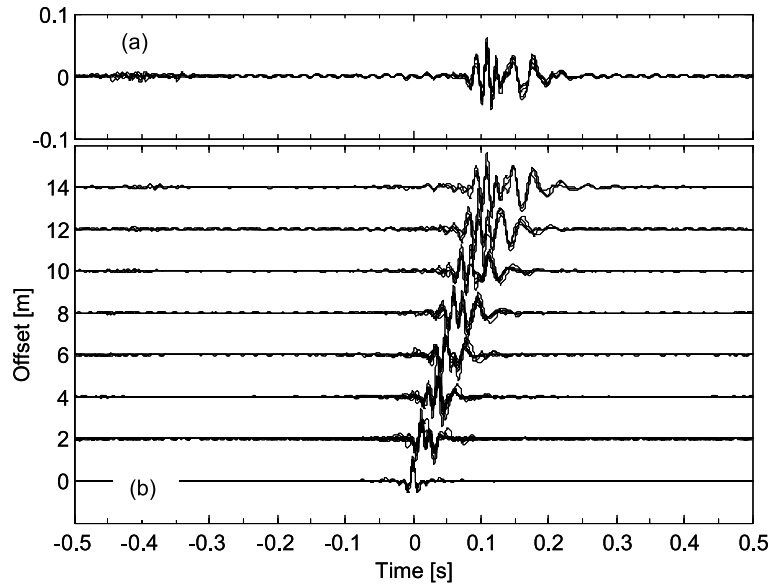


Figure 4.2: (a) Superposition of five 1-min-long correlations between an accelerometer pair separated by 14 m. (b) Superposition of seismic sections obtained from correlation between accelerometer pairs separated by the same range. The signal is normalized at each range.

The advantage of the correlation process is to perform an average on the different sources by stacking them with the appropriate time lag without the need for any synchronization between the source and the receivers. The superposition of the correlations of 1-min long signals is thus just a verification of the repeatability of the steps. Stacking the five correlations of 1-min-long records is equivalent to directly correlating a 5-min-long signal.

To obtain a seismic section from the correlation process, each signal recorded at one receiver is correlated with the signal recorded by accelerometer 1 or 8 located at each extremity of the line array. The seismic section clearly shows wave propagation along the geophone array [Fig. 4.2(b)]. Taking geophone 1 or 8 as the reference signal does not modify the seismic section. This shows that seismic propagation from left to right is identical to propagation from right to left on the 14-m-long seismic array. The medium can then be assumed as one dimensional (1D) in the frequency bandwidth of the recordings.

The 1D argument can be pushed even further. Each receiver pair spaced by the same range should give the same correlation function as the propagation does not depend on the receiver locations but only on the separation between receivers. For each range, there are several equivalent receiver pairs. Figure 4.2(b) shows the superposition of the time-domain correlations of the equivalent pairs. For shorter ranges, there are more equivalent pairs (7 for 2 m, 6 for 4 m, etc.), which results in an efficient stacking and a better SNR. Figure 4.3(a) shows the seismic section obtained after all averaging operations. After a 10-min total recording, SNR is above 30 dB. Both phase dispersion and geometrical spreading during propagation are retrieved.

The final seismic section clearly reveals the presence of two surface waves, with mean group velocities of about 90 and 120 m/s. Those low group velocities are good indications of two surface waves identified as Rayleigh modes.

We insist on the fact that this section was obtained from 10 min of unsynchronized human steps only, which makes it nearly a passive method. To get the same result with classical active seismic

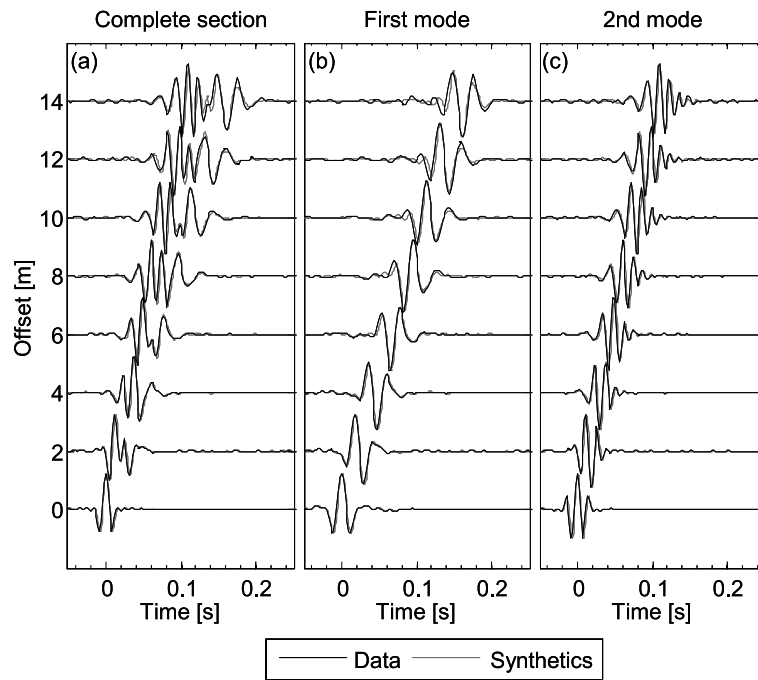


Figure 4.3: (a) Seismic section obtained from the correlation process after all averaging operations. The signal-to-noise ratio is above 30 dB for each trace. Phase dispersion is clearly observed for two surface waves with group velocity of about 90 and 120 m/ s. (b), (c) The two propagating modes have been extracted from the F–K plot in Fig. 4.4. Dispersion of each mode is clearly observed. Synthetics obtained from the velocity model in Fig. 4.5(a) show a good agreement with experimental data.

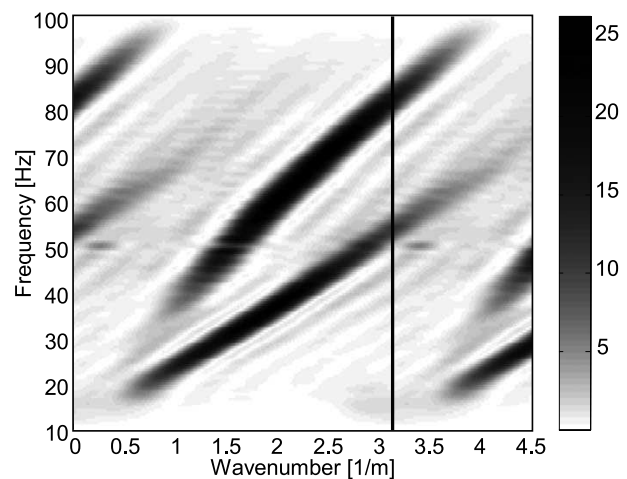
techniques, much more time would have been needed to synchronize numerous sledgehammer blows. The “passive” method presented here is thus (1) easy to implement, as there is a large flexibility in the array configuration, (2) fast, as it takes only about 30 min to complete the array deployment and the recording, and (3) simple, as there is no synchronization task and processing is performed in real time.

► 4.1.4 Geophysics inversion

Given the high SNR of the seismic section obtained from the correlation process, we completed the analysis by a classical surface wave seismic inversion [Shapiro et al., 2005, Wathelet et al., 2004].

A frequency–wave-number (F–K) transform was applied to the seismic section [Fig. 4.3(a)] to measure phase velocity dispersion curves [Fig. 4.4]. As the geophones are evenly spaced by a 2-m distance, the largest wave number satisfying the aliasing criterion is $2\pi/d = \pi$. In Fig. 4.4, this corresponds to the black vertical line. Higher k ’s are wrapped, and appear as low wave numbers. In this simple case, the wave-number spectrum can be extended by unwrapping the k axis.

Figure 4.4: Frequency–wavenumber (F–K) transform of the seismic section obtained in Fig. 4.3(a). The largest measurable wave number according to the aliasing criterion is $2\pi/d$ with $d=2$ m (black vertical line). The aliasing in the F–K diagram is resolved by unwrapping the k axis. The shape of the two modes on the F–K diagram reveals dispersive modes.



In the F–K diagram, modes are separated and extracted by masking a region of the F–K plot and returning to the range/time domain [Fig. 4.3(b) and 4.3(c)]. This operation cannot be performed directly in the range/time domain since the two modes are mixed at short ranges.

From Fig. 4.4, phase velocity curves are measured for each mode separately using the relation $k = 2\pi f/c$, where k is the wavenumber, f is the frequency, and c is the phase velocity [Fig. 4.5(b)]. A Monte Carlo inversion is performed simultaneously on the two modes to obtain the most-likely shear velocity profile in the medium [Wathelet et al., 2004] [Fig. 4.5(a)]. The minimum frequency of the surface waves being larger than 20 Hz for a maximal wavelength of 12 m, the model is not constrained for depth deeper than 6 m. Similarly, the P-velocity profile and the density were poorly constrained by the surface wave dispersion curves and arbitrarily set to 900 m/s and 1.5 g/cm³, respectively.

To improve confidence in the model, synthetic sections are computed separately for the two Rayleigh modes [Fig. 4.3(b) and 4.3(c)] [Herrmann, 1987]. The computed traces are very close to the extracted modes. To check the relative weight of each mode, their sum is compared to

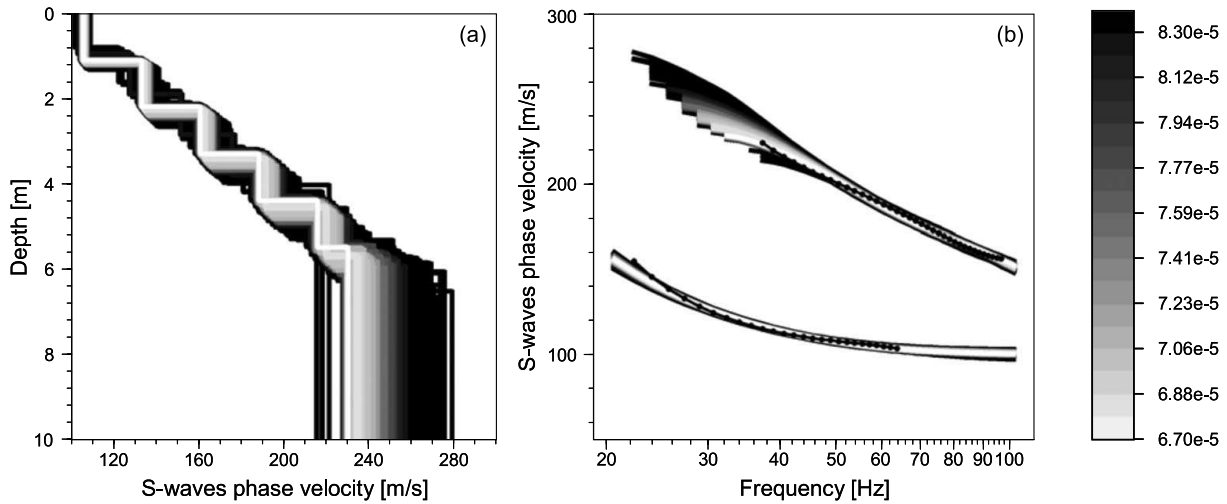


Figure 4.5: (a) Distribution of shear velocity profiles obtained from a Monte Carlo inversion using the dispersion curves of the two surface-wave modes. (b) Distribution of phase dispersion curves for the two surface modes in the Monte Carlo inversion. Black spots correspond to the experimental phase-velocity curves obtained from Fig. 4.4. Colorbar scales correspond to a misfit increase of 24%.

the original seismic section before mode separation [Fig. 4.3(a)]. As the complete seismic section is well retrieved, the relative weights of the computed and experimental modes are identical, which confirms the quality of the inversion model. To increase the precision of the model, a full wave inversion could be performed that would benefit from the high SNR of the seismic section [Fig. 4.3(a)]. In this case, the inversion kernel should be modified to include both amplitude and time dispersion information.

► 4.1.5 Conclusion

A new method to retrieve surface waves at small scales has been developed. This method, which combines passive processing and active sources is simpler, faster, and easier to implement than classical active seismic techniques. It is also versatile and can be easily adapted to various bottom configurations and expected wavelengths.

A classical seismic inversion was performed using the observed seismic section. The high SNR allows us to get an accurate model for shear velocity in the (0–7 m) depth.

Future works will investigate the case of two-dimensional (2D) media. In this case the stacking process performed in a 1D case won't be possible, but differences between intercorrelations at the same range but at different places could provide a 2D (range + depth) inversion result.

► Acknowledgment

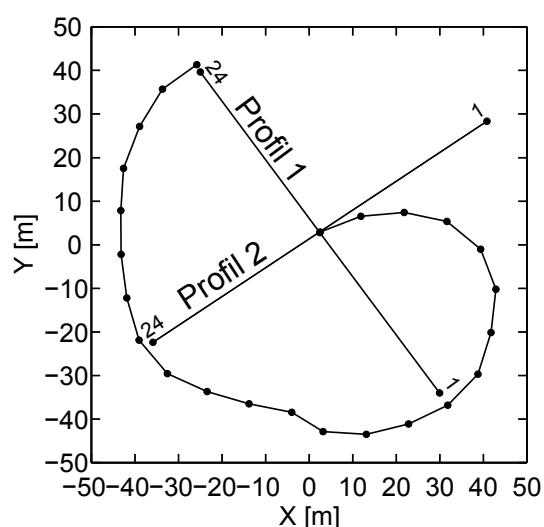
The authors are grateful to Marc Wathelet for his advice in the use of the Monte Carlo inversion software GEOPSY.

4.2 Étude d'un milieu latéralement inhomogène

Nous avons voulu appliquer la méthode décrite dans le paragraphe précédent à un milieu latéralement inhomogène. Cette étude se fera en deux étapes. Dans la première les variations du milieu sous les réseaux utilisés ne seront pas prises en compte. Ce sera l'occasion de comparer les résultats obtenus par corrélation à ceux obtenus grâce à des enregistrements actifs (coups de masse synchronisés), les traitements habituellement appliqués à ces derniers faisant l'hypothèse du milieu homogène. La deuxième approche consiste à profiter du grand nombre de trajets reconstruits par corrélation et de les traiter indépendamment en prenant en compte les variations locales de vitesse par un algorithme de tomographie.

Des enregistrements de pas humains et de coup de masse synchronisés ont donc été effectués au même endroit, et dans les mêmes conditions expérimentales, afin de pouvoir comparer les résultats obtenus. La zone étudiée est un champ de maïs, dans sa période d'hivernage, le déploiement ayant été effectuée au mois de novembre 2007. Les premières dizaines de centimètres du sol ont donc été fortement remaniées. Les coups de masse synchronisés ont été donnés à l'extrémité de deux profils linéaires perpendiculaires (voir la figure 4.6), avec 24 géophones verticaux et horizontaux (installés en position transverse) séparés de 4 m. Des pas non synchronisés ont été enregistrés à l'aide du même dispositif, également avec des géophones verticaux et horizontaux, ainsi que dans une configuration en spirale comprenant 24 géophones verticaux présentée sur la figure 4.6. L'inter-trace du réseau en spirale est d'environ 10 m. Les signaux de pas ont été générés par plusieurs personnes marchant en même temps dans une zone d'un mètre de diamètre environ, constituant un point source. Ces points sources sont situés de chaque côté des profils linéaires, dans leur alignement, ainsi qu'autour de chaque géophone du réseau en spirale. Les positions de ces points source ont été choisies de façon à pouvoir exploiter le plus de trajets inter-récepteurs possible. La durée d'enregistrement des pas est de 4 minutes pour les profils linéaires et de 15 secondes autour de chaque géophone pour la spirale.

FIG. 4.6 – Les deux profils linéaires et le réseau en spirale utilisés. Chacun des 3 déploiements comporte 24 géophones. Les numéros sur les deux profils linéaires indiquent les numéros des géophones, qui définissent l'orientation des axes (1→24) pour l'analyse de ces profils. Des acquisitions actives et passives ont été réalisées pour les deux profils linéaires, avec des géophones verticaux et horizontaux, les points source étant situés de part et d'autre des profils. Seuls des enregistrements de pas ont été réalisés pour la spirale, avec des géophones verticaux, les points source étant situés tour à tour au niveau de chacun des récepteurs.



► 4.2.1 Étude des profils linéaires

Hypothèse du milieu homogène

Les expériences actives ont été traitées selon la méthode MASW [Park et al., 1999]. Pour chaque point source, 5 coups de masse synchronisés ont été enregistrés, puis les traces ont été sommées afin d'augmenter le rapport signal sur bruit. Les coups de masse sont donnés verticalement sur un plaque de frappe pour les profils utilisant des géophones verticaux, et horizontalement sur une poutre lestée pour ceux utilisant des géophones horizontaux. Des exemples de sections sismiques ainsi obtenues sont présentés sur les figures 4.7a et 4.7b. A partir de ces sections actives, une analyse FK est réalisée afin de mesurer les courbes de dispersion des ondes dans le milieu. Cette analyse suppose que le milieu de propagation est le même sous tous les éléments du réseau, puisqu'elle est basée sur la décomposition en ondes planes du champ d'onde enregistré (voir le paragraphe 2.3.1).

Les enregistrements de pas ont été corrélés afin de retrouver les fonctions de Green entre toutes les paires de géophones des profils. L'utilisation de géophones horizontaux ou verticaux ne change ni la façon dont sont générés les signaux (*i.e.*, les pas), ni le traitement appliqué aux enregistrements. Les deux acquisitions sont cependant traitées indépendamment, les géophones verticaux permettant de reconstruire les ondes de Rayleigh tandis que les horizontaux apportent des informations sur les ondes de Love. Le milieu n'étant pas homogène, les fonctions de Green reconstruites par corrélation dépendent de la distance entre les récepteurs, mais aussi de leur position. L'hypothèse du milieu homogène est tout de même utilisée dans un premier temps afin de retrouver un milieu homogène équivalent. Les fonctions de Green reconstruites pour des géophones éloignés de la même distance, mais situés en des endroits différents du profil, sont ainsi considérées comme identiques et les fonctions de corrélation moyennées, comme décrit dans l'expérience précédente (voir la figure 4.2). Les figures 4.7c et 4.7d présentent les résultats ainsi obtenus pour les géophones verticaux et horizontaux traités séparément. Ces sections sismiques sont à comparer à celles obtenues de façon active à partir des coups de masse et présentées sur la même figure. Les modes de propagation des sections actives et passives sont les mêmes, mais avec des amplitudes relatives différentes. Ces différences d'amplitude ont pour origine la différence d'excitation entre les pas et la masse avec plaque de frappe, amplifiée par le processus de corrélation. En effet l'opération de corrélation étant mathématiquement une multiplication de spectres, les différences d'énergies relative entre les modes sont mises au carré et les spectres rendus plus piqués. La figure 4.8 montre les spectres des sections actives et passives, moyennés sur toutes les traces de chacune des sections. Les pas ont un spectre d'excitation plus basse fréquence que les coups de masse, ce qui est visible dans les séries temporelles de la figure 4.7. Les courbes de dispersion des ondes de Rayleigh et de Love sont mesurées à partir des sections active et passive et présentées sur la figure 4.9. Le grand rapport signal sur bruit des sections de corrélation permet la mesure de la dispersion dans la même gamme de fréquences que les sections actives malgré la différence de spectres. Les résultats obtenus à partir des sections corrélées sont très proches de ceux obtenus avec la méthode MASW classiquement utilisée à ces échelles dans les études géophysiques. Les courbes de dispersion de la figure 4.9 sont celles d'un milieu homogène équivalent.

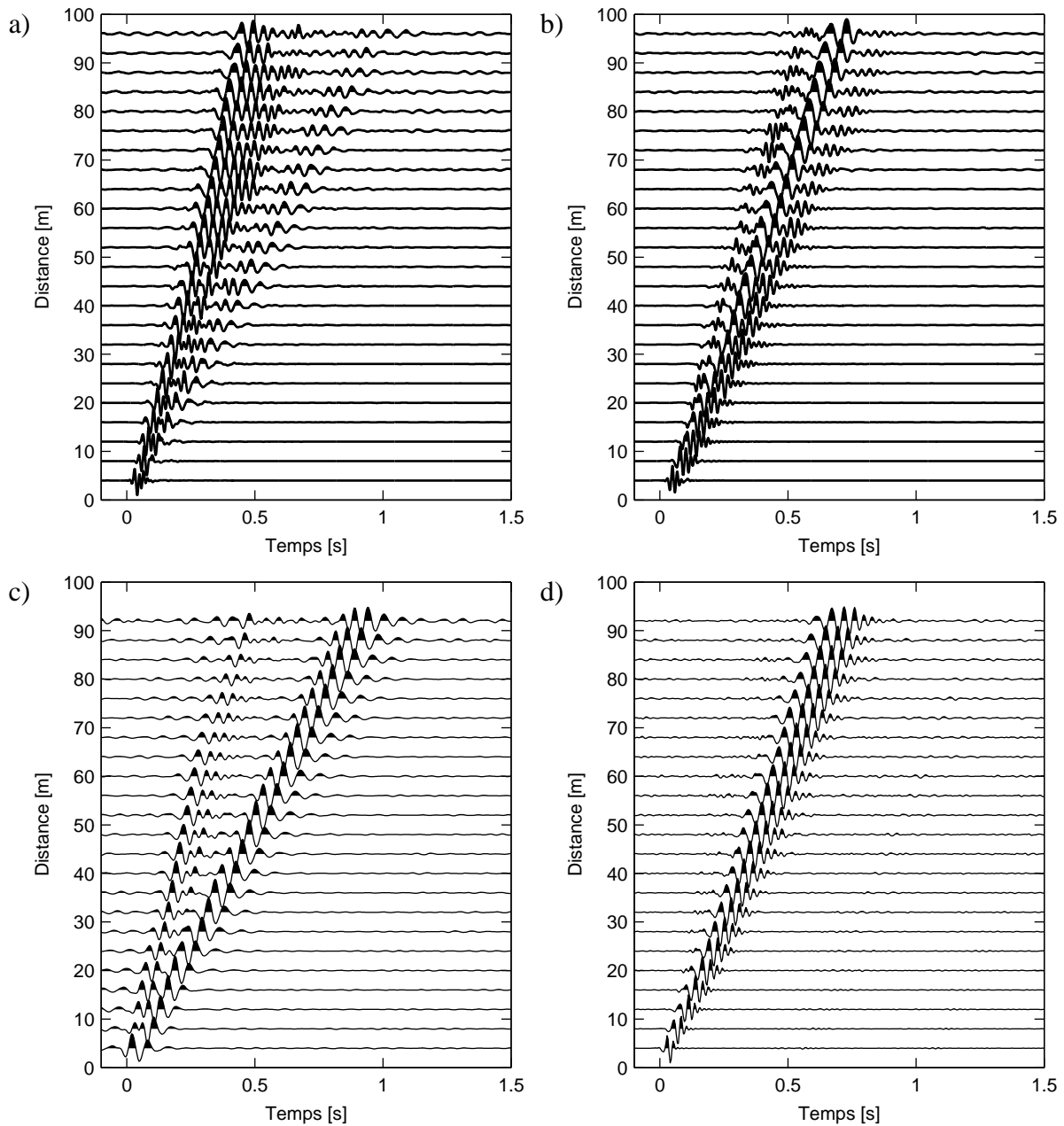


FIG. 4.7 – Sections sismiques pour le premier profil : **a.** enregistrements actifs, géophones verticaux ; **b.** enregistrements actifs, géophones horizontaux ; **c.** enregistrements passifs corrélés, géophones verticaux ; **d.** enregistrements passifs corrélés, géophones horizontaux. Les traces sont normalisées en amplitude pour la représentation. Les sections actives et passives présentent les mêmes modes de propagation, mais avec des amplitudes relatives et des contenus fréquentiels différents.

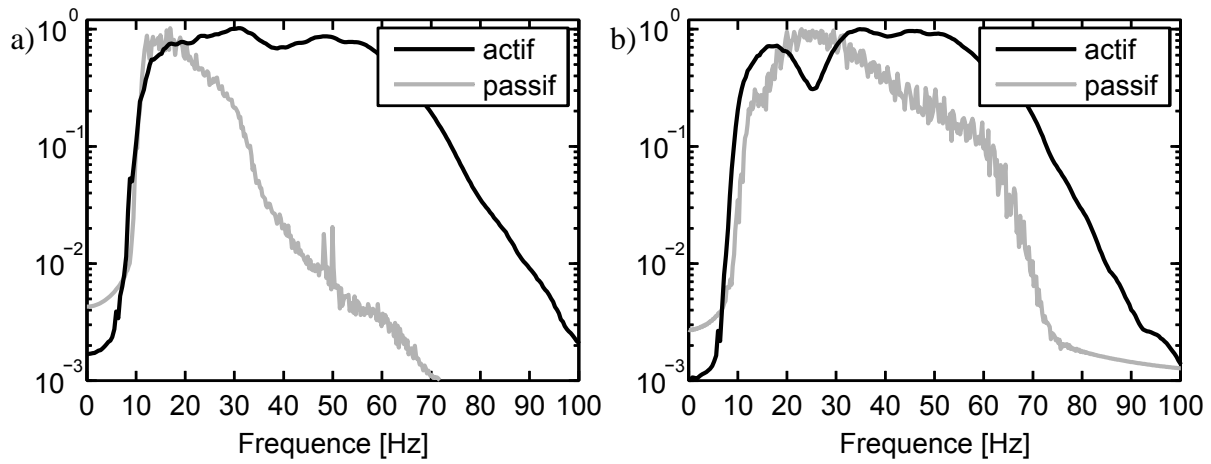


FIG. 4.8 – Spectres normalisés moyens des sections passives et actives de la figure 4.7, pour les géophones **a.** verticaux et **b.** horizontaux. Les sections corrélées sont plus basses fréquence que les sections actives.

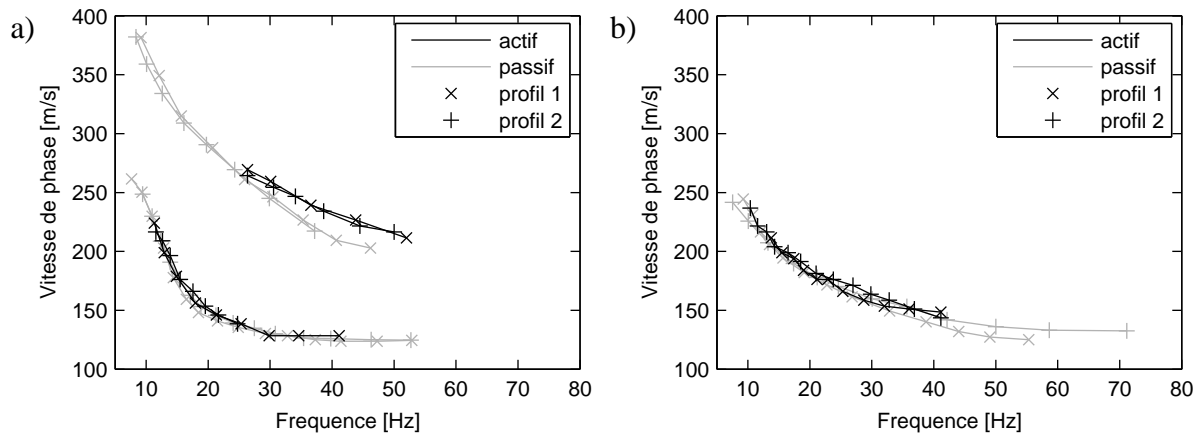


FIG. 4.9 – Courbes de dispersion obtenues à partir des sections actives et passives, pour les deux profils en déplacement **a.** vertical (Rayleigh) et **b.** horizontal (Love). Les deux méthodes donnent des résultats proches, mais dans des bandes de fréquences différentes. Le mode harmonique étant très peu énergétique dans la section de corrélation en déplacement vertical (figure 4.7c), la mesure de sa courbe de dispersion est peu précise.

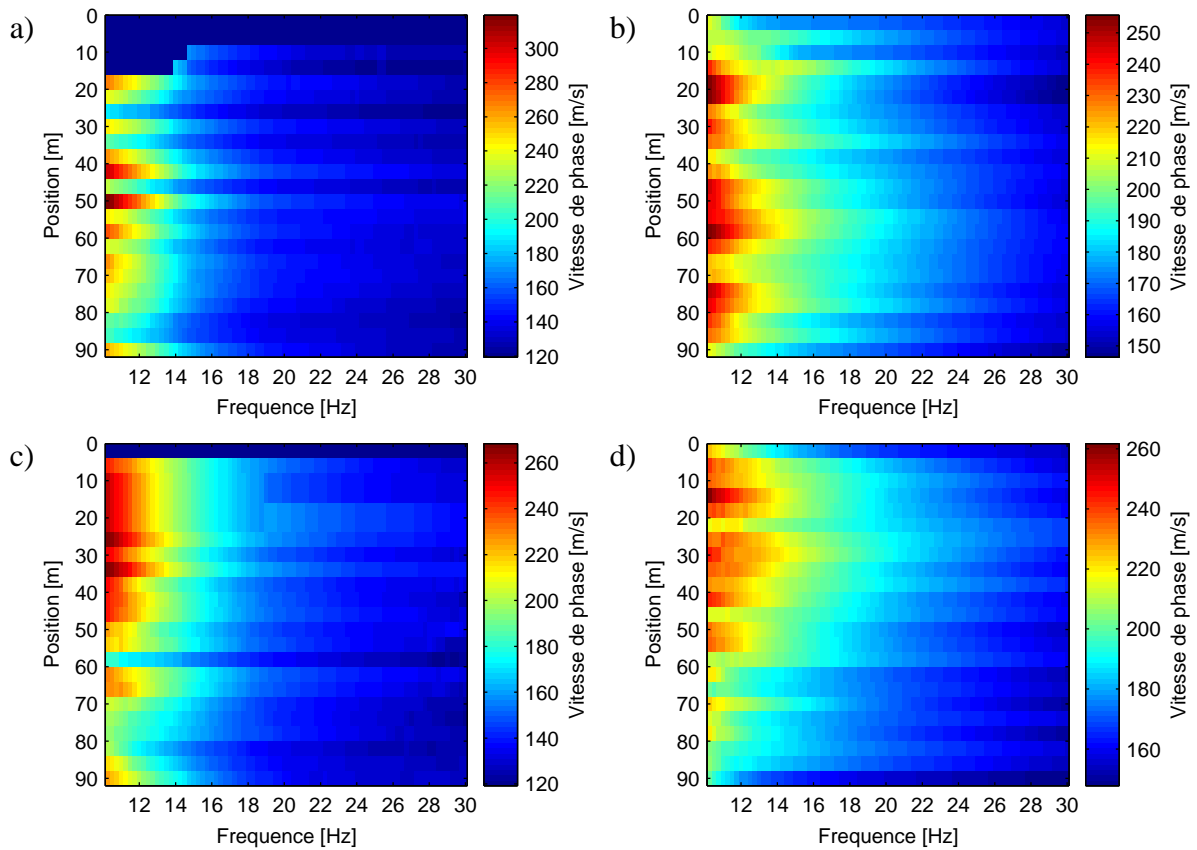


FIG. 4.10 – Courbes de dispersion en fonction de la position le long des profils pour **a.** le profil 1 et les géophones verticaux, **b.** le profil 1 et les géophones horizontaux, **c.** le profil 2 et les géophones verticaux, **d.** le profil 2 et les géophones horizontaux. Les zones bleu foncé des figures **a.** et **b.** correspondent à des positions et des fréquences pour lesquelles la mesure des temps de propagation n'est pas assez précise pour réaliser l'inversion. L'évolution de la vitesse moyenne le long des profils est cohérente avec les courbes de dispersion mesurées dans l'approximation de milieu homogène présentée sur la figure 4.9.

Prise en compte des variations du milieu le long des profils

L'hypothèse de milieu homogène faite précédemment permet de mesurer la dispersion dans un milieu équivalent. Cette hypothèse n'est cependant pas valable pour le milieu étudié ici. J'ai donc modifié le traitement présenté jusqu'ici afin de prendre en compte d'éventuelles variations du milieu le long des deux profils : les fonctions de corrélation sont calculées entre toutes les paires de géophones disponibles, mais sans être moyennées pour les récepteurs séparés par la même distance comme précédemment. Pour chaque paire de récepteurs, les temps de parcours des ondes de Rayleigh et de Love (suivant le type de géophone considéré) sont mesurés en fonction de la fréquence. Un algorithme d'inversion (décrit au paragraphe 2.4) est ensuite utilisé afin de trouver des modèles de dispersion expliquant les temps mesurés. Cette approche n'est possible qu'à partir des données corrélées, car elle nécessite la connaissance des temps de parcours pour tous les couples de récepteurs. Pour pouvoir l'utiliser sur les données actives, il aurait fallu frapper des coups de masse au niveau de chacun des 24 géophones des profils. Les modèles de vitesse de phase des ondes de Rayleigh et de Love obtenus pour les deux profils sont présentés sur la figure 4.10. Chaque ligne de ces

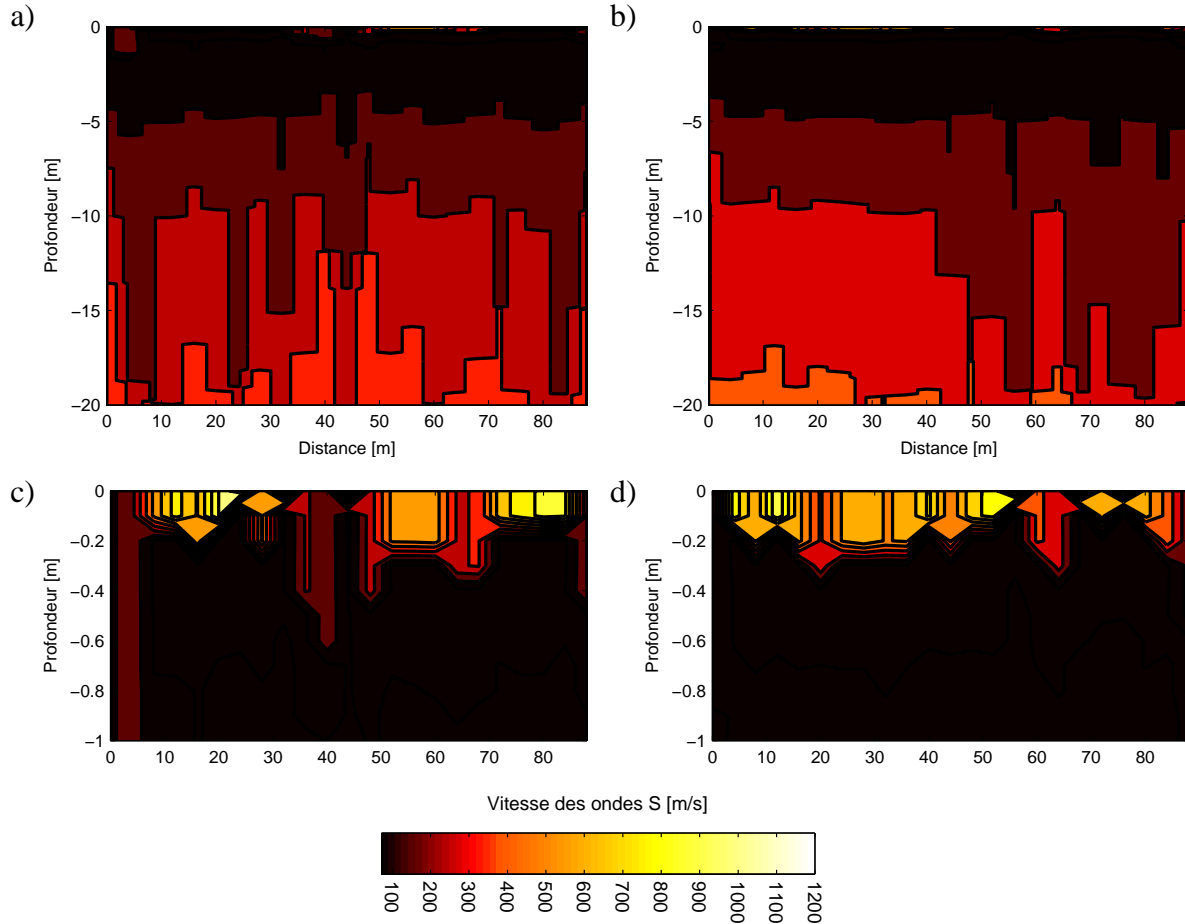


FIG. 4.11 – Modèles de vitesse des ondes S sous **a.** le profil 1, et **b.** le profil 2. Les figures **c.** et **d.** sont des agrandissements de ces modèles de vitesse pour les faibles profondeurs. Les modèles obtenus semblent montrer un pendage orienté dans le sens du second profil. La couche fine à haute vitesse en surface est bien contrainte dans l'inversion grâce à l'utilisation conjointe des dispersions des ondes de Rayleigh et de Love.

figures peut être interprétée comme une courbe de dispersion pour le milieu situé à la verticale de la position correspondante le long du profil. Ces figures montrent l'évolution latérale de la dispersion dans le milieu. Les deux profils se croisant en leur centre, les mesures à la position $X = 46$ m de chacun des profils constituent deux mesures indépendantes pour le même milieu.

Modèles de vitesse sous les profils et étude de sensibilité

L'inversion de ces courbes de dispersion pour les ondes de Rayleigh et de Love, en chacun des points de mesure, est réalisée afin d'obtenir les profils de vitesse 2D (ou pseudo-2D, la tomographie étant réalisée indépendamment pour chaque point de mesure [Luo et al., 2008]) des ondes S en fonction de la profondeur, présentés sur la figure 4.11. Les inversion ont été réalisées en utilisant un algorithme exploratoire de type *Monte Carlo* [suite logicielle *Sesarray*, Wathelet, 2008]. Ces modèles de vitesse montrent des variations qui font penser à un pendage orienté dans le sens du second profil. En effet le modèle de vitesse varie peu le long du premier profil, contrairement à ce que l'on obtient le long du second profil où il y a une différence entre les modèles obtenus sous

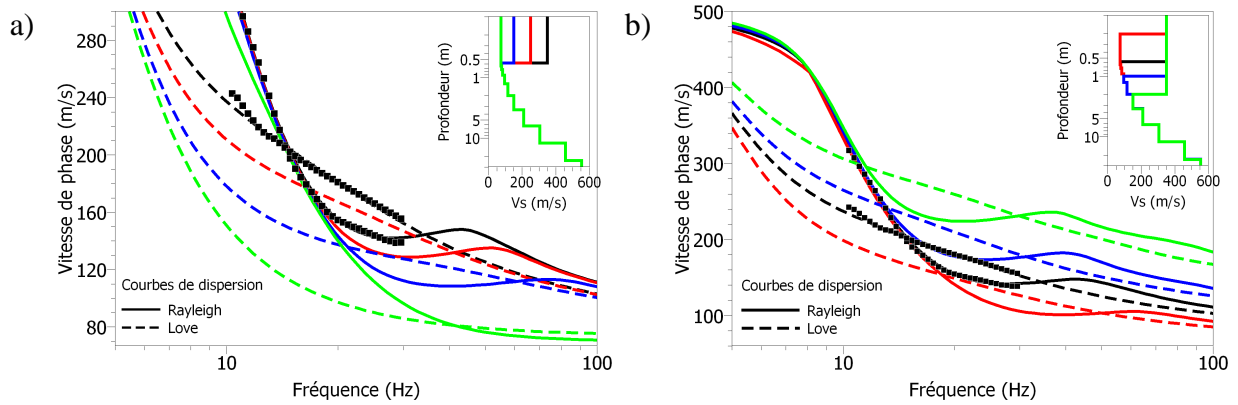


FIG. 4.12 – Étude de sensibilité des courbes de dispersion à **a.** la vitesse et **b.** l'épaisseur d'une couche fine en surface. Les courbes de dispersion sont calculées pour les modes fondamentaux de Rayleigh et de Love dans différents modèles de vitesse présentant une couche fine à haute vitesse en surface. Les carrés noirs indiquent les courbes de dispersion mesurées pour le second profil à l'abscisse $X = 50$ m (figure 4.9), et le modèle tracé en noir est le modèle correspondant de la figure 4.11b. Le point de croisement des courbes de dispersion de Rayleigh et de Love contraint l'épaisseur et la vitesse de la couche de surface.

chaque moitié du profil. Cela peut s'expliquer par le fait que le premier profil est perpendiculaire à la direction du pendage. Les deux profils se croisent en leur milieu, et les modèles obtenus à la verticale de ce point sont comparables, alors qu'ils constituent des inversions indépendantes. Les modèles obtenus présentent une couche très fine à très haute vitesse en surface.

L'épaisseur de cette couche étant faible devant les longueurs d'ondes mises en jeu (de l'ordre de 5 mètres à 30 Hz), nous avons effectué une étude de sensibilité présentée sur la figure 4.12 afin de justifier la résolution de cette couche. Cette étude de sensibilité a consisté à calculer les courbes de dispersion des modes fondamentaux de Rayleigh et de Love dans différents modèles présentant une couche fine en surface, dont on a fait varier l'épaisseur et la vitesse. Les résultats obtenus indiquent que le point de croisement des courbes de dispersion de Rayleigh et de Love varie lorsque l'un au moins de ces deux paramètres varie. La mesure de ce point de croisement contraint donc bien la vitesse et l'épaisseur de la couche de surface.

La résolution de cette couche de surface peut paraître en contradiction avec le critère de sensibilité communément admis qui stipule que le mode fondamental des ondes de Rayleigh est sensible à la vitesse dans le milieu à une profondeur d'environ un tiers de la longueur d'onde. Ce critère est obtenu en considérant la forme de ce mode en fonction de la profondeur. Il faut cependant rappeler que ce critère empirique est lié à la profondeur maximale à laquelle le mode est sensible. En effet, les ondes de surface sont définies par les pôles d'une équation décrivant l'interaction entre les ondes de volume au niveau de la surface libre. Ces pôles étant très sensibles aux conditions limites à cette interface, l'ajout d'une couche fine en surface peut fortement perturber ces ondes, même si les ondes de volume ne seront que très peu affectées. Cette sensibilité des ondes de surface aux conditions de surface libre est problématique dans le sens où une faible erreur de mesure sur la dispersion de ces ondes peut engendrer des erreurs importantes sur le modèle obtenu après inversion. L'utilisation conjointe des ondes de Rayleigh et de Love permet cependant de stabiliser l'inversion

et la rend plus robuste aux erreurs de mesure.

Ayant montré que la résolution de la couche en surface était liée aux informations combinées des modes fondamentaux de Rayleigh et de Love, et non à la connaissance de chacune de ces courbes à haute fréquence, nous avons réalisé 3 inversions en supprimant de plus en plus d'informations sur la dispersion à hautes fréquences. La première inversion est ainsi réalisée avec une courbe de dispersion mesurée entre 10 et 50 Hz, la deuxième entre 10 et 30 Hz, et la dernière entre 10 et 20 Hz. Afin de ne pas introduire de biais lié à d'éventuelles erreurs dans la détermination des courbes de dispersion, les mesures utilisées sont issues de courbes théoriques calculées à partir du modèle présenté en noir sur les graphiques de droite de la figure 4.13. Cette figure présente en grisé un ensemble de modèles issus de l'inversion (colonne de droite), ainsi que les courbes de dispersion associées (colonne de gauche). Pour les trois inversions, la dispersion des modèles permettant d'expliquer la partie des courbes de dispersion utilisée est faible, ce qui indique que l'inversion est bien contrainte. Tous les modèles obtenus présentent la couche de surface du modèle théorique. Les courbes de dispersion calculées à partir du résultat de la dernière inversion (utilisant en entrées des mesures jusqu'à 20 Hz) présentent pour les fréquences supérieures à 20 Hz, non utilisées pour l'inversion, les caractéristiques de la courbe théorique (représentée dans bande 10-50 Hz par les points noirs sur la figure de la première inversion).

► 4.2.2 Utilisation du réseau en spirale

Les enregistrements de pas localisés autour d'un récepteur permettent de retrouver la fonction de Green entre ce récepteur et chacun des autres géophones du réseau. Une tomographie faite à partir de la mesure des temps de propagation entre tous les récepteurs du réseau permet alors de retrouver une image de la vitesse des ondes S en fonction de la profondeur dans le milieu. Afin que la tomographie soit performante, il faut que les temps de propagation mesurés soient triés afin de minimiser les erreurs de mesure. Pour cela, j'ai profité du fait qu'on dispose de deux mesures indépendantes pour chaque trajet mesuré, correspondant à des points sources situés autour de chacun des deux récepteurs de la paire considérée. La cohérence entre les deux temps de propagation donne un critère de sélection des mesures.

Les mesures de temps de propagation sont effectuées à chaque fréquence grâce à la phase de la transformée de Fourier des signaux corrélés. Une fois toutes les mesures effectuées, les temps de propagation entre tous les couples sont triés selon le critère évoqué plus haut et les mesures retenues sont utilisées pour former l'image de la vitesse dans le milieu à chaque fréquence. La zone d'environ 120 m de diamètre est ainsi totalement couverte à partir d'un seul déploiement de 24 géophones.

La figure 4.14 montre les cartes de courbes de dispersion mesurées pour les deux premiers modes de Rayleigh. Pour chaque position de cette figure, les courbes de dispersion sont inversées afin de retrouver le modèle de vitesse des ondes S en fonction de la profondeur pour la zone, en fonction de la position. Sur les mesures de vitesse des ondes de Rayleigh pour le mode harmonique, présentées sur la figure 4.14b, des variations de vitesses de phase très locales apparaissent, ce qui ne semblent pas correspondre à des variations réelles du milieu mais plutôt à des erreurs de mesure lors du pointé des temps de parcours pour ce mode. Ce mode ne sera donc pas utilisée pour l'inversion tomographique, celle-ci pouvant être très sensible aux erreurs de mesure. Les mesures de courbes de

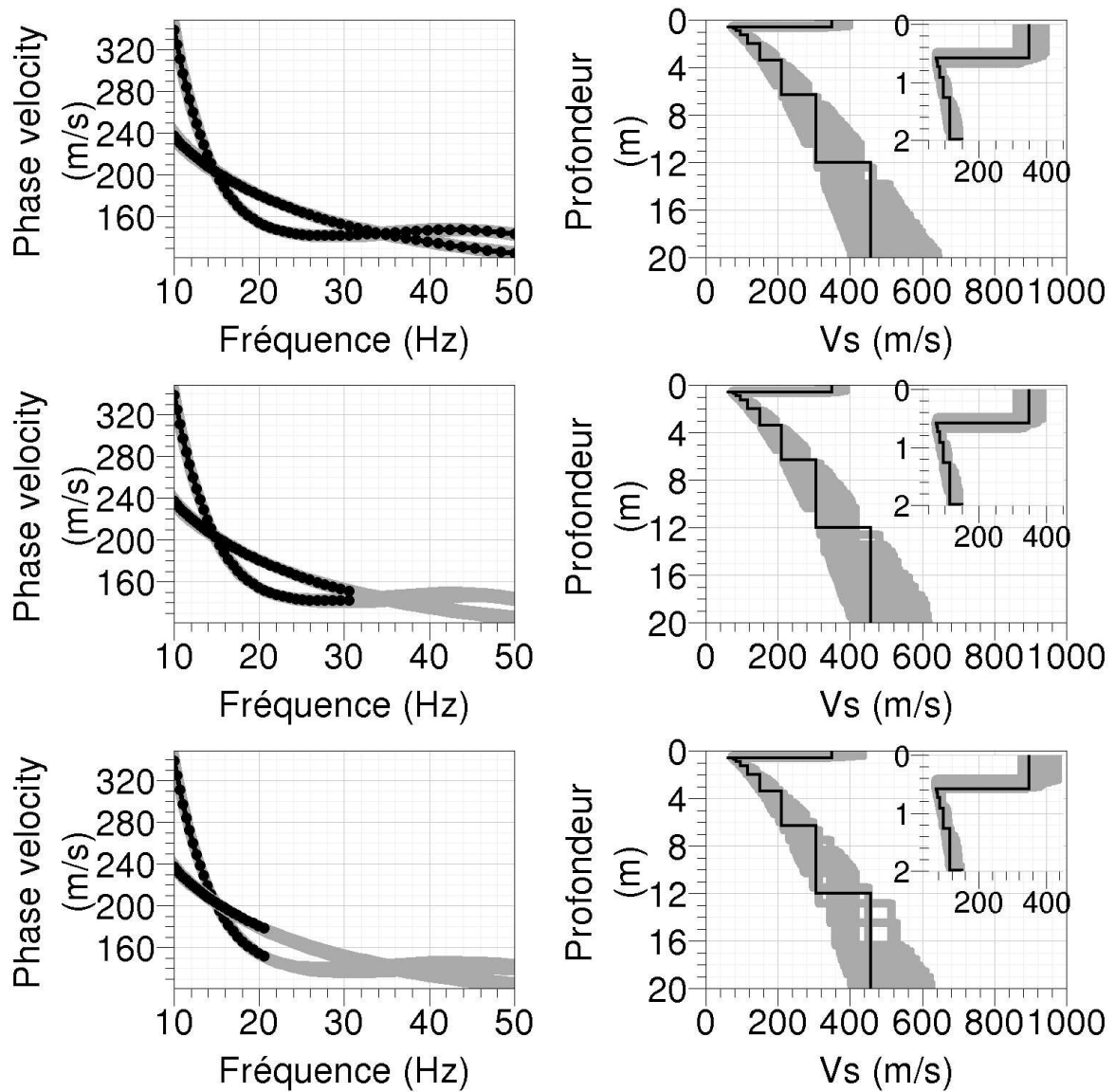


FIG. 4.13 – Trois inversions ont été effectuées à partir de la même courbe de dispersion théorique, calculée à partir du modèle indiqué en noir sur chacune des figures de la colonne de droite, et renseignée dans le processus d'inversion dans des bandes de fréquence de plus en plus étroites : 10-50 Hz ; 10-30 Hz ; 10-20 Hz. Les points noirs sur les graphiques de la colonne de gauche indiquent les mesures utilisées pour chacune des 3 inversions. Les inversions utilisent un algorithme exploratoire de type *Monte Carlo*. Les parties grisées indiquent, pour les graphiques de droite, un ensemble de modèles expliquant les courbes de dispersion, et l'ensemble des courbes de dispersion associées sur les graphiques de gauche.

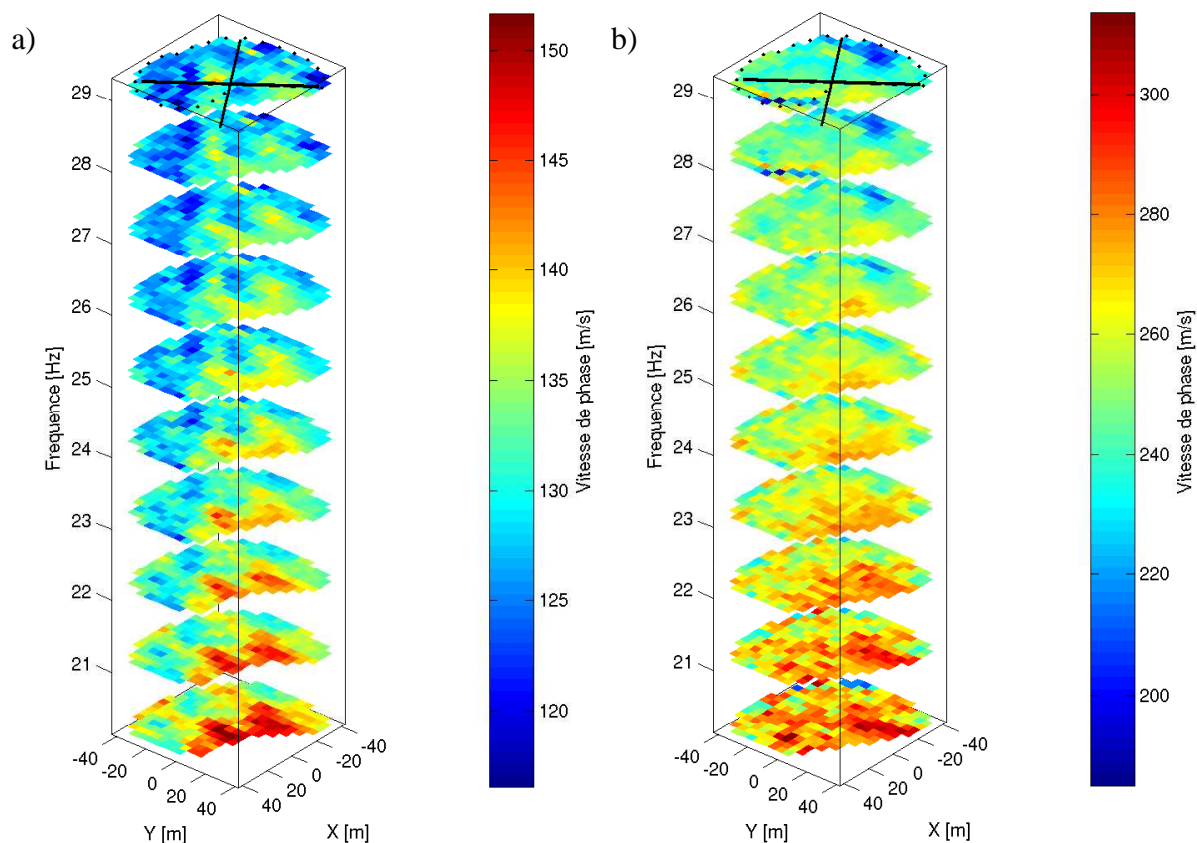
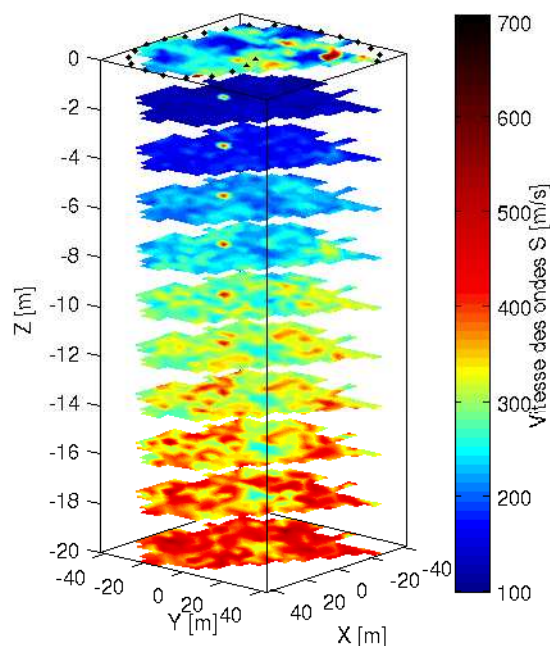


FIG. 4.14 – Modèles de vitesse des ondes de Rayleigh en fonction de la position et de la fréquence pour **a.** le mode fondamental et **b.** le premier mode harmonique. Les variations très locales font penser à des erreurs dans la mesure des temps de propagation et ne semblent pas physiques. Les deux lignes indiquent les positions des profils linéaires.

FIG. 4.15 – Modèle de vitesse pseudo-3D des ondes S en fonction de la profondeur, obtenu à partir de l'inversion tomographique des courbes de dispersion du mode fondamental de Rayleigh présentées sur la figure 4.14a. Ce modèle présente une inversion de vitesse à faible profondeur, déjà obtenue à partir de l'inversion des profils linéaire. Le pendage évoqué plus haut lors de l'analyse des profils linéaires n'apparaît pas sur cette figure. Cela peut être dû à une mauvaise contrainte de l'inversion due à un manque d'information, les mesures sur les ondes de Love n'étant pas disponibles.



dispersion présentées ici sont cohérente avec celles obtenues à partir des profils linéaires présentés sur la figure 4.10. On retrouve en particulier le changement de vitesse de phase pour le mode fondamental de Rayleigh présenté sur la figure 4.10c pour le second profil.

Le modèle de vitesse pseudo-3D ainsi obtenu est présenté sur la figure 4.15. Ce modèle est pseudo-3D et non purement 3D dans le sens où il résulte de la combinaison d'inversions pour chaque pixel de l'image finale, faisant chacune l'hypothèse d'un milieu stratifié. Une inversion purement 3D serait effectuée en prenant en compte en un seul processus toutes les informations sur les variations latérales des courbes de dispersion. Le modèle obtenu ne présente pas les variations latérales de la profondeur de l'interface à environ 10 mètres de profondeur que l'on voit sur les inversions des profils linéaires de la figure 4.11, que l'on a interprété comme un pendage. Cette différence peut avoir plusieurs origines :

1. une erreur d'inversion due à une contrainte moins forte sur les modèles. En effet seuls des géophones verticaux ont été utilisés pour le profil en spirale, ne permettant ainsi qu'une mesure de la dispersion des ondes de Rayleigh. Lors de l'inversion des profils linéaires, une mesure sur le mode fondamental de était également utilisée, et nous avons vu que l'utilisation des informations conjointes sur ces deux ondes permet de fortement contraindre l'inversion.
2. une mauvaise précision dans la détermination des courbes de dispersion due à des erreurs de pointé des temps de propagation ou à un lissage des variations spatiales de ces courbes lié à la méthode de mesure. Nous avons déjà écarté les mesures sur le mode harmonique de Rayleigh pour cette raison. Les même erreurs peuvent être présentes dans les mesures relatives au mode fondamental, même si cela ne semble pas apparaître sur la figure 4.14a.

On retrouve néanmoins sur la figure 4.15 les caractéristiques principales des modèles obtenus lors de l'étude des profils linéaires, et en particulier l'inversion de vitesse à faible profondeur. Les valeurs de la vitesse en fonction de la profondeur sont également compatibles avec celles de la figure 4.11 obtenues pour les profils linéaires.

► 4.2.3 Conclusion sur cette étude

L'utilisation d'une méthode de tomographie à des fins d'imagerie sur des profils linéaires nous a permis de prendre en compte les variations du modèle le long de ces profils, alors que l'hypothèse de milieu homogène est habituellement utilisée pour les études à cette échelle. Cette méthode nécessite de connaître individuellement les temps de propagation entre tous les éléments du réseau, ce qui est rendu possible à moindre coût par les techniques de corrélation. De plus, l'utilisation combinée des mesures relatives aux ondes de Rayleigh et de Love permet de fortement contraindre l'inversion et d'avoir ainsi une grande précision sur le modèle de proche surface.

A partir du réseau en spirale, nous avons pu obtenir une image pseudo-3D d'une zone de 120 m de diamètre environ, en utilisant seulement 24 points de mesure. Pour cette étude, les imprécisions sur la mesure des temps de propagation et l'absence de mesure du déplacement horizontal transverse empêche d'obtenir une aussi bonne précision sur le modèle final que celle obtenue avec les profils linéaires. On retrouve néanmoins les caractéristiques principales du modèle obtenu dans l'étude des profils linéaires.

Il faut rappeler ici que les méthodes d'inversion de courbes de dispersion permettent de retrouver un modèle expliquant les courbes de dispersions, mais qu'il y a non unicité du résultat. Ainsi, les modèles présentés sur les figures 4.11 et 4.15 sont les modèles trouvés présentant des courbes de dispersion les plus proches de celles mesurées, mais rien n'assure qu'ils correspondent au modèle réel. La précision de ces modèles est difficile à évaluer. On peut, dans le meilleur des cas, présenter une classe de modèles ayant des écarts comparables à la courbe de dispersion mesurée.

4.3 Étude de la Solfatare, Italie

► 4.3.1 Présentation de la Solfatare

La Solfatare (*Solfatara* en italien), située dans la ville de Pozzuoli à proximité de Naples en Italie, est l'un des volcans de la caldera des Champs Phlégréens. Le système magmatique des Champs Phlégréens est encore actif, puisque la dernière éruption remonte à 1538 au Monte Nuovo. L'activité volcanique de la Solfatare se traduit principalement par des fumerolles et des sources chaudes soufrées. Les Champs Phlégréens sont sujet aux phénomènes de bradysisme, c'est-à-dire que la région dans son ensemble subit des phases d'élévation et d'effondrement. Durant la dernière crise sismique des années 1982–1984, une élévation de 1,8 m centrée sur la ville de Pozzuoli a été mesurée. Lors de cette crise, plus de 16 000 séismes ont été recensés, à des profondeurs de 0 à 4 km, majoritairement localisés dans la Solfatare (surtout en ce qui concerne les plus énergétiques). La déformation de la zone de la Solfatare est le résultat de variations de la pression des fluides dans le système géothermal sous-jacent [Bonafede & Mazzanti, 1998]. Les remontées de gaz et la température élevée des fluides créent un panache hydrothermal ascendant dont le sommet se trouve au milieu du cratère de la Solfatare, et est très visible sur l'image présentée figure 4.16. Cette zone centrale est appelée *Fangaia* en raison de la nature boueuse du sol et comprend plusieurs mares d'eau bouillonnantes. La différence de niveau de la nappe phréatique entre l'intérieur et l'extérieur du cratère est de quelques dizaines de mètres.

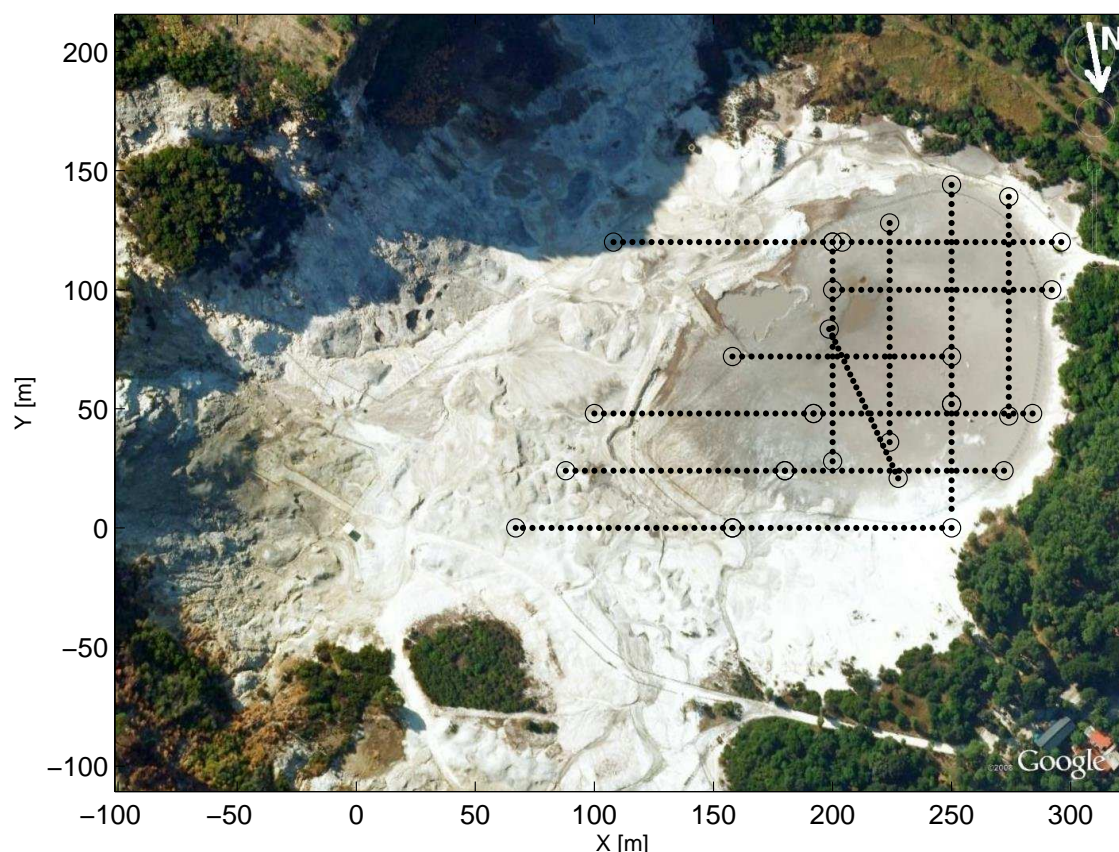


FIG. 4.16 – Vue aérienne de la Solfatare avec les différents déploiements. Les cercles indiquent les extrémités des 16 profils de 24 géophones. Les réseaux ont été déployés dans la *Fangaia* où la nappe phréatique est affleurante. Les taches situées autour des points de coordonnées ($X = 175$ m, $Y = 90$ m) et ($X = 210$ m, $Y = 90$ m) sont des mares de fluide en ébullition. Les principales fumerolles se situent sur la gauche de l'image.

L'objectif de l'étude est de caractériser la structure du panache hydrothermal, et en particulier de reconnaître les différentes phases du fluide (liquide ou vapeur) grâce à différentes méthodes géophysiques : tomographie de résistivité, tomographie de vitesse de surface et cartographie thermique. Je m'intéresserai ici principalement aux mesures sismiques, et donc à la tomographie de vitesse de surface. Une comparaison entre les résultats obtenus à partir des différentes méthodes sera toutefois effectuée à la fin de ce paragraphe.

Cette zone présente de fortes variations latérales, tant au niveau des propriétés du milieu (profondeur de la nappe, porosité...) que des sources locales de bruit. Nous avons déployé au centre du cratère une ligne de 24 géophones verticaux de fréquence de coupure 4,5 Hz, que nous avons déplacés, définissant ainsi les 16 profils indiqués sur la figure 4.16. Pour chaque profil, nous avons enregistré des séries de sources non synchronisées, durant 30 secondes à chaque fois (soit une soixantaine de sources par acquisition), localisées à chaque extrémités du profil ainsi qu'entre chaque géophones (soit 25 acquisitions par profil), ce qui constitue les données « actives ». Nous avons également effectué des acquisitions de bruit ambiant de 4 minutes pour chaque profil, constituant les données « passives ». Le traitement des données passives sera exposé dans ce manuscrit au paragraphe 5.3, dans un chapitre consacré à la localisation des sources de bruit.

► 4.3.2 Étude des données actives

Pour chaque profil de 24 géophones, nous avons enregistré 25 séries de coups de masse non synchronisés, de 30 secondes chacune, localisés entre chaque paire de récepteurs ainsi qu'à chaque extrémité du profil, dans son alignement. La distance entre les géophones est de 4 mètres, et la plus courte distance source-récepteur est donc de 2 m. L'environnement de la Solfatare étant naturellement très bruyé, les enregistrements ne sont pas toujours dominés par les coups de masse générés, en particulier pour des récepteurs situés loin du point source. Lorsque cela n'est pas le cas, les fonctions de corrélation sont dominées par un pic correspondant à une (ou plusieurs) source dont on ne connaît pas la position. Elles ne sont alors pas utilisables pour une mesure de temps de parcours car la fonction de corrélation n'a *a priori* rien à voir avec la fonction de Green entre les récepteurs, comme nous l'avons vu au chapitre 2. Une sélection soignée des enregistrements et des fonctions de corrélation calculées est donc essentielle. Pour chaque paire de récepteurs, le traitement retenu est le suivant :

- les points source situés à moins de 4 mètres de l'un des deux récepteurs ne sont pas considérés afin d'éviter les problèmes de saturation ou de réponse non linéaire des géophones, ainsi que les effets de champ proche. Ceux situés entre les deux récepteurs de la paire sont également écartés, puisqu'ils ne permettent pas de reconstruire la fonction de Green entre les récepteurs.
- pour chaque point source restant, les enregistrements sont filtrés entre 8 et 20 Hz et la fonction de corrélation normalisée entre les 2 géophones est calculée. Seules les fonctions de corrélation ayant un rapport signal sur bruit (défini ici comme le rapport entre le maximum de la fonction de corrélation et l'écart type du bruit au temps longs) supérieur à 10 sont retenues.
- étant donné que les points source sont traités indépendamment, on s'attend à n'obtenir dans la fonction de corrélation qu'un seul pic, dont la position (temps positifs ou négatifs) est connue. Seules les fonctions de corrélation ayant le maximum du « bon côté » sont retenues.
- si après cette sélection il reste au moins 3 fonctions de corrélation, elles sont moyennées pour donner la fonction de Green reconstruite pour la paire considérée. Si il en reste moins de 3, le couple est abandonné.
- si le couple est conservé, on mesure le maximum de l'enveloppe de la fonction de Green. Le temps du maximum donne le temps de parcours du paquet d'onde entre les éléments du couple considéré, associé à la vitesse de groupe dans le milieu.

Cette sélection drastique des fonctions de corrélation ne permet de retenir qu'une petite fraction des temps de parcours mesurables (*i.e.*, des 24×23 paires possibles au sein d'un même profil). En particulier, les mesures entre les couples éloignés de plus de 20 mètres sont rarement sélectionnées, étant donné le faible rapport signal sur bruit des corrélations pour ces couples. Le grand nombre de mesures disponibles permet cependant d'en garder suffisamment pour pouvoir retrouver la vitesse autour de tous les profils.

Une fois les temps de parcours mesurés pour tous les couples retenus de tous les profils, la deuxième étape consiste à réaliser l'inversion proprement dite. L'algorithme utilisé est celui décrit au paragraphe 2.4, appliqué à chaque profil. Un lissage du modèle est introduit dans l'inversion via une longueur de corrélation utilisée dans la définition de la covariance sur le modèle (voir l'équation (2.14)), fixée à $d_0 = 16$ m. Le modèle de vitesse de groupe des ondes de Rayleigh est présenté sur la figure 4.17, en haut. La cohérence des mesures présentées sur cette figure est remarquable, les mesures effectuées sur chacun des profils étant totalement indépendantes et non

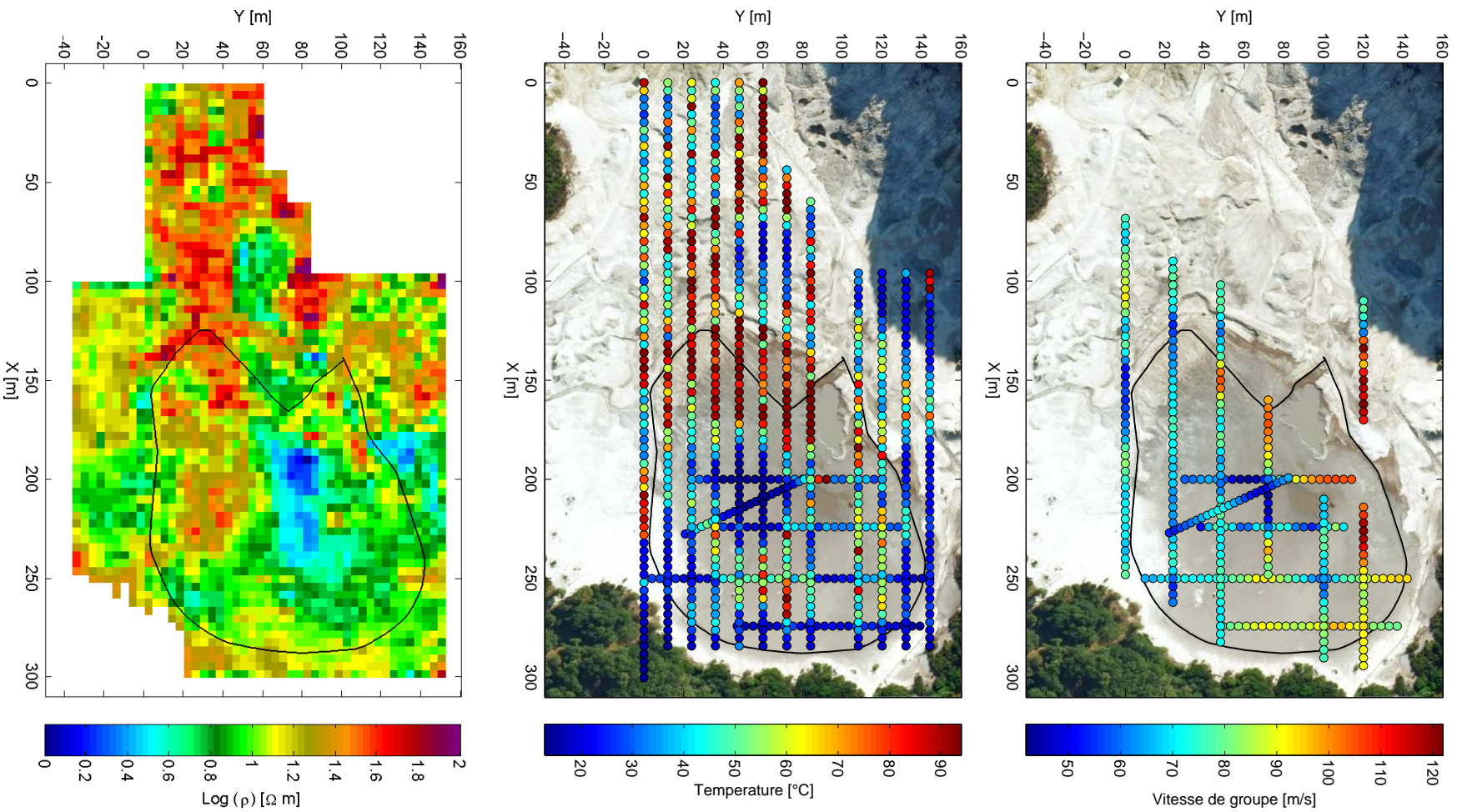


FIG. 4.17: Les différentes mesures effectuées dans le cratère de la Solfatare : mesures de vitesse de groupe des ondes de surface (obtenues par tomographie), mesure de température et résistivité de surface.

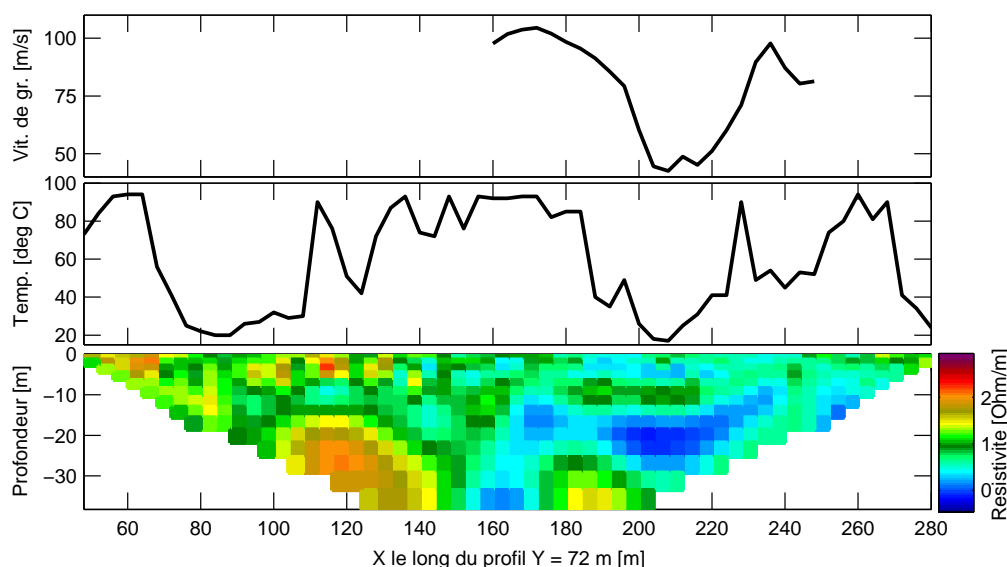


FIG. 4.18 – Vitesse de groupe, température et modèle de résistivité inversé le long du profil Y = 72 m. Les zones de surface conductrices humides correspondent à des zones de faible température. Lorsque la température augmente, le fluide se vaporise et la vapeur chaude assèche le milieu, le rendant moins conducteur et plus rapide. En profondeur le fluide reste sous forme liquide, sa température d'ébullition étant plus importante à cause de la pression.

simultanées. Cette carte montre des zones à faible vitesse, principalement au centre de la *Fangaia*, et des zones à forte vitesse comme dans le coin supérieur gauche de la figure.

La figure 4.17 présente également les autres mesures effectuées dans cette zone. Les données de température ont été obtenues à l'aide d'un thermocouple placé en chaque point de mesure à une dizaine de centimètres de profondeur. Celles de résistivité résultent d'une acquisition selon un protocole dipôle-dipôle, la carte présentée ici étant la résistivité de surface. Il y a une correspondance entre les zones de forte vitesse et les zones de haute température et de forte résistivité. On peut interpréter cette correspondance par un assèchement du sol dans les zones de haute température par évaporation du fluide. Le milieu asséché présente alors une vitesse des ondes de Rayleigh plus importante et une augmentation de résistivité par rapport aux zones plus humides. Les observations sur place montrent que ces zones ont un aspect de surface croûté, avec en quelques endroits des cavités situées à quelques centimètres de profondeur. Les quelques centimètres au dessus de la cavité sont alors très durs et secs. Sur la figure 4.17, les zones de faible résistivité ne coïncident pas avec les mares visibles en surface. L'absence de zone conductrices au niveau des mares peut être expliquée par l'absence de mesure de résistivité le long du profil Y = 96 m, précisément à cause de la présence de ces mares. La profondeur de ces mares étant faibles (estimée à deux mètres), la faible résistivité des mares n'est pas mesurée par les profils situés autour, à une distance trop importante pour y être sensible.

La figure 4.18 montre les différentes mesures extraites le long d'un profil défini par Y = 72 m dans les coordonnées de la figure 4.17. Sur ces profils, on voit bien le lien entre la vitesse de groupe des ondes de Rayleigh et la température. Sur le profil inversé en résistivité, les zones bleues conductrices montrent une zone liquide en profondeur, à la verticale de laquelle la vitesse est faible

et la température basse. Dans le modèle proposé ici, ce fluide a une température élevée, supérieure à 100°C, mais reste sous forme liquide en profondeur à cause de la pression qui élève la température d'ébullition. Lorsque ce fluide remonte en surface, il rentre en ébullition à cause de la diminution de pression et passe sous forme vapeur et devient moins conducteur, excepté en quelques points en surface où sont situées les mares, visibles sur la figure 4.17. En dehors de ces mares, le passage de la vapeur assèche le milieu qui devient moins conducteur. Les zones de sortie de vapeur sont caractérisées par une forte température.

Afin de pouvoir réaliser une cartographie de la vitesse des ondes de Rayleigh à plus grande échelle et en 2 dimensions, comme pour l'étude présentée précédemment avec le réseau en spirale, une nouvelle campagne d'acquisition va être réalisée fin 2008. Cette campagne utilisera un matériel permettant des déploiements de réseaux avec des ouvertures plus grandes, et surtout une couverture simultanée de tout le cratère grâce à environ 80 points de mesure. Afin de limiter l'atténuation, les capteurs permettront une acquisition plus basse fréquence. Les sources utilisées ne seront pas des coups de masse, dont le spectre d'excitation est trop haute fréquence et l'énergie trop faible pour enregistrer le signal au delà de quelques dizaines de mètres, mais un pot vibrant permettant l'utilisation de *sweeps* afin de transmettre plus d'énergie dans le sol.

L'analyse des données de bruit ambiant enregistrées sur les profils est présentée à la fin de ce manuscrit, au paragraphe 5.3. Elle a pour objectif de localiser les sources de bruit afin de mieux comprendre le fonctionnement du système hydrothermal sous-jacent.

4.4 Conclusion sur l'utilisation de sources non synchronisées

Les expérimentations présentées ici valident l'utilisation de sources non synchronisées. Elle permet de s'affranchir de la distribution des sources de bruit ambiant tout en gardant un système d'acquisition simple ne nécessitant pas de déclenchement synchronisé. L'utilisation d'une distribution de source adaptée permet l'étude d'un milieu inhomogène, selon des profils linéaires ou en 3 dimensions à partir d'un réseau déployé en spirale. Le faible nombre de capteurs utilisés est alors un avantage, tant en ce qui concerne le coût financier de l'acquisition qu'au niveau de sa simplicité.

Ces études ont été l'occasion de mettre en évidence la grande sensibilité des ondes de Rayleigh et de Love à la vitesse des ondes dans les premiers mètres sous le réseau. Cette sensibilité, qui peut être pénalisante à cause de l'importance que prennent les erreurs de mesure, peut permettre une très bonne résolution de la proche surface quand l'inversion est bien contrainte, comme c'est le cas lors d'une mesure conjointe de la dispersion des ondes de Rayleigh et de Love.

Cette approche est particulièrement utile lorsque le milieu est très bruyé, avec de nombreuses sources de bruit locales, comme c'est le cas dans le cratère de la Solfatare. L'étude complexe de la distribution de ces sources n'est alors pas nécessaire. La simplicité du dispositif permet également de multiplier les mesures et de combattre ainsi leur faible rapport signal sur bruit.

Les sources utilisées dans le cadre de cette thèse sont des pas humains ou des coups de masse, choisies pour leur simplicité de mise en œuvre, mais n'importe quelle source peut être utilisée, afin

d'adapter le spectre d'excitation et l'énergie transmise dans le sol aux besoins de l'étude. Ainsi, il l'utilisation de pots vibrants devrait permettre des mesures à plus basses fréquences, et avec des distances de propagation plus importantes.

Bibliographie chapitre 4

Bonafede, M. & Mazzanti, M., 1998. Modelling gravity variations consistent with ground deformation in the Campi Flegrei caldera (Italy), *Journal of Volcanology and Geothermal Research*, **81**, 137–157.

Bruno, P. P. G., Ricciardi, G. P., Petrillo, Z., Di Fiore, V., Troiano, A., & Chiodini, G., 2007. Geophysical and hydrogeological experiments from a shallow hydrothermal system at Solfatara Volcano, Campi Flegrei, Italy: response to caldera unrest, *Journal of Geophysical Research (Solid Earth)*, **112**(B6), 6201.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**, 547–549.

Gouédard, P., Roux, P., & Campillo, M., 2008. Small-scale seismic inversion using surface waves extracted from noise cross correlation, *The Journal of the Acoustical Society of America, Express Letters*, **123**(3), EL26–EL31.

Herrmann, R. B., 1987. Computer programs in seismology, Department of Earth and Atmospheric Sciences, Saint Louis University, Saint Louis, available for free at <http://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html>.

Luo, Y., Xia, J., Liu, J., Xu, Y., & Liu, Q., 2008. Generation of a pseudo-2d shear-wave velocity section by inversion of a series of 1d dispersion curves, *Journal of Applied Geophysics*, **64**(3-4), 115–1124.

Park, C. B., Miller, R. D., & Xia, J., 1999. Multichannel analysis of surface waves, *Geophysics*, **64**(3), 800–808, - MASW.

Roux, P. & Kuperman, W. A., 2004. Extracting coherent wavefronts from acoustic ambient noise in the ocean, *The Journal of the Acoustical Society of America*, **116**(4), 1995–2003.

Roux, P., Sabra, K. G., Gerstoft, P., & Kuperman, W. A., 2005. P-waves from cross-correlation of seismic noise, *Geophysical Research Letters*, **32**, L19303.

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Physical Review E*, **69**(4), 046610.

Wapenaar, K., 2003. Synthesis of an inhomogeneous medium from its acoustic transmission response, *Geophysics*, **68**(5), 1756–1759.

Wathelet, M., 2008. An improved neighborhood algorithm: parameter conditions and dynamic scaling, *Geophysical Research Letters*, **35**, L09301, - a citer quand on utilise dinver (avec/ou un autre de 2004).

Wathelet, M., Jongmans, D., & Ohrnberger, M., 2004. Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surface Geophysics*, **2**, 211–221, - a citer pour qd on utilise dinver (avec/ou un autre papier de Marc de 2008).

Caractérisation du bruit ambiant

Sommaire

5.1 Méthodes	134
5.1.1 Formation de voie (<i>Beamforming</i>)	134
5.1.2 Décomposition en valeurs singulières	136
5.1.3 A propos du rang de la matrice de densité spectrale croisée	137
5.1.4 Ambiguïté du modèle de vitesse	138
5.2 Étude de la vallée de Waimangu, Nouvelle-Zélande	138
5.2.1 Présentation de la vallée de Waimangu	138
5.2.2 Influence du cycle	140
5.2.3 Traitement fréquentiel	144
5.2.4 Comparaison avec la polarisation spontanée	144
5.3 Étude passive de la Solfatare, Italie	147
Bibliographie chapitre 5	152

J'ai présenté jusqu'ici les méthodes de reconstruction de fonctions de Green à l'aide de corrélations, en adaptant le traitement aux propriétés du champ d'onde enregistré. Nous avons vu que l'isotropie du champ joue un rôle particulier, et que quand cette hypothèse est mise en défaut il faut adapter le traitement des fonctions de corrélations en faisant une hypothèse supplémentaire, celle du milieu homogène. Il est donc essentiel de pouvoir déterminer, à partir d'enregistrements de bruit ambiant sur un réseau, les propriétés d'un champ d'onde afin de mettre en place une méthode de corrélation adaptée. Ce chapitre sera consacré à cette étude, en utilisant essentiellement la formation de voie. Après un paragraphe d'introduction à cette méthode, deux études expérimentales seront proposées sur des sites en Nouvelle-Zélande et en Italie. Ces applications se limitent à l'étude de la distribution de sources, et le problème de la reconstruction de fonctions de Green en prenant en compte cette distribution ne sera pas abordé.

5.1 Méthodes

► 5.1.1 Formation de voie (*Beamforming*)

La *formation de voie* (*beamforming* en anglais) est une technique qui consiste à appliquer des retards (positifs ou négatifs) aux éléments d'un réseau (en émission ou en réception) en fonction de leur positions relatives, selon une loi de phase décrivant un front d'onde prédéterminé. Par exemple, si on veut générer une onde plane à partir d'un réseau de sources, il faut appliquer à chaque source un retard de façon à ce qu'elle contribue de façon cohérente au front d'onde plan créé par les autres sources du réseau. La loi de retard à utiliser dépend de la vitesse des ondes dans le milieu. De la même façon, si on considère un réseau de récepteurs, et qu'on veut enregistrer préférentiellement les ondes planes provenant d'une direction donnée, on peut appliquer à chacun des enregistrements un retard adapté, puis sommer les enregistrements, afin que les ondes planes provenant effectivement de cette direction s'ajoutent de façon cohérente (NB : cette méthode est appelée en géophysique *frequency-wavenumber analysis* car la loi de phase dépend du vecteur d'onde et de la fréquence – voir les paragraphes 2.3.1 et 3.1.2). Lorsque les fronts d'onde recherchés ne sont pas plans (*i.e.*, que les sources ne sont pas en champ lointain), la loi de phase change. On trouve alors cette méthode dans la littérature sous le nom de *matched field processing* [Baggeroer et al., 1988, 1993, Kuperman and Turek, 1997]. Ce traitement peut être décrit selon un autre point de vue : appliquer une loi de phase, puis sommer les traces, est équivalent à considérer un jeu de données synthétiques (appelé *replica*), correspondant à ce qu'on enregistrerait à chacun des capteurs si la source était à une position donnée dans un modèle de source considéré (champ proche ou lointain par exemple), et de le comparer au jeu de données réelles. J'ai utilisé la formation de voie en réception seulement, afin de retrouver, dans un milieu bruité, la position de la (ou des) source de bruit à partir d'un réseau de capteurs. Pour cela, différents replicas sont calculés, correspondant à différentes positions \vec{r}_s possibles de la source. La comparaison entre ces replicas et les données effectivement enregistrées donne la réponse $P(\vec{r}_s)$ de la formation de voie pour la position \vec{r}_s . $P(\vec{r}_s)$ peut être considérée comme la probabilité de présence de la source à la position \vec{r}_s , même si nous verrons plus loin que ce n'est pas à proprement parler une probabilité.

Le traitement s'effectue dans le domaine de Fourier. On note $d_i(\omega)$ le champ enregistré sur le récepteur i pour la pulsation ω . La matrice de densité spectrale croisée $K(\omega) = (K_{ij}(\omega))$ est d'abord calculée :

$$K_{ij}(\omega) = d_i(\omega) d_j^*(\omega) \quad , \quad (5.1)$$

où l'étoile représente le transposé hermitien. Cette matrice n'est rien d'autre que la corrélation entre les champs enregistrés à chaque capteur, pour la pulsation ω . Si on note N_r le nombre de récepteurs dans le réseau, $K(\omega)$ est une matrice de taille $N_r \times N_r$. Cette matrice contient tous les auto-spectres et les inter-spectres, et inclue ainsi toutes les relations entre les enregistrements, en fonction de la fréquence.

Si on considère un milieu homogène de vitesse c , le replica pour une source située en $\vec{r}_s = (x_s, y_s, z_s)$ s'écrit

$$e_i(\vec{r}_s, \omega) = \frac{1}{4\pi r_i^s} e^{i\omega(t-r_i^s/c)} \quad , \quad (5.2)$$

où r_i^s représente la distance entre la position \vec{r}_s de la source et le récepteur i . La comparaison avec les données peut se faire de différentes façons. La plus simple est de considérer le coefficient de corrélation entre les données $d = (d_i)$ et le replica $e = (e_i)$. La réponse de cette méthode, appelée méthode de Bartlett, s'écrit alors

$$P(\vec{r}_s, \omega) = e^*(\vec{r}_s, \omega) K(\omega) e(\vec{r}_s, \omega) \quad . \quad (5.3)$$

P n'est pas à proprement parler une probabilité de présence, car elle n'est pas normalisée. En effet, si $e(\vec{r}_s, \omega)$ est exactement identique aux données, on a bien une amplitude maximale pour P , égale à 1. Cependant, l'amplitude de la source réelle n'est pas prise en compte dans la formule, et lorsque celle-ci n'est pas égale à 1, ou lorsque la source n'est pas impulsionnelle en temps et/ou en espace (ce qui est le cas en pratique puisque cette méthode est généralement appliquée à des enregistrements de bruit ambiant), l'amplitude absolue de P n'a pas de signification. Seules les variations relatives de P peuvent donc être interprétées.

La méthode de Bartlett peut être raffinée en utilisant une comparaison non linéaire entre le replica et les données. L'avantage des méthodes non linéaires est que la décroissance de la réponse $P(\vec{r}_s)$ lorsque \vec{r}_s s'éloigne de la position réelle de la source est plus importante que dans la méthode de Bartlett. La méthode *Minimum Variance Distortionless Response* (MVDR, [Capon, 1969]) est une méthode purement non linéaire. Son principe est de considérer un replica w_{MV} calculé à partir de e , mais possédant une amplitude jouant le rôle de pondération pour chaque récepteur i dans le calcul de P . La valeur de la pondération est calculée de telle façon que la norme de $w_{MV}^* K w_{MV}$ soit minimale, tout en imposant $w_{MV}^* e = 1$. La solution de cette optimisation est

$$w_{MV}(\vec{r}_s, \omega) = \frac{K^{-1}(\omega) e(\vec{r}_s, \omega)}{e^*(\vec{r}_s, \omega) K^{-1}(\omega) e(\vec{r}_s, \omega)} \quad . \quad (5.4)$$

La probabilité de présence P de la source est alors

$$P_{MV}(\vec{r}_s, \omega) = w_{MV}^*(\vec{r}_s, \omega) K(\omega) w_{MV}(\vec{r}_s, \omega) \quad (5.5)$$

$$= \frac{1}{e^*(\vec{r}_s, \omega) K^{-1}(\omega) e(\vec{r}_s, \omega)} \quad . \quad (5.6)$$

En contrepartie de sa grande résolution, Cox et al. [1987] ont montré que la sensibilité au bruit incohérent (*i.e.*, n'étant pas lié à une position de source, comme le bruit électronique d'acquisition par exemple) de cette méthode est très fortement accrue. Ils ont alors proposé une méthode plus robuste, appelée *White Noise Constraint* (WNC), qui permet de contrôler le degré de non linéarité utilisé grâce à un paramètre δ compris entre 0 et 1. Dans cette méthode, le replica s'écrit

$$w_{WN}(\vec{r}_s, \omega) = \frac{(K(\omega) + \varepsilon(\vec{r}_s, \omega) I)^{-1} e(\vec{r}_s, \omega)}{e^*(\vec{r}_s, \omega) (K(\omega) + \varepsilon(\vec{r}_s, \omega) I)^{-1} e(\vec{r}_s, \omega)} \quad , \quad (5.7)$$

où I est la matrice identité de taille $N_r \times N_r$ et $\varepsilon(\vec{r}_s, \omega)$ est un paramètre auto adaptatif qui s'ajuste de façon à satisfaire $w_{WN}^* w_{WN} \leq \delta^{-2}$. Ce replica vérifie également $w_{WN}^* e = 1$. Quand le paramètre ε tend vers l'infini, $w_{WN} \approx e$ et la méthode est linéaire. Inversement, $\varepsilon = 0$ permet de retrouver l'algorithme purement non linéaire MVDR. La relation entre ε et δ est complexe. On peut néanmoins

retenir que $\delta = 1$ engendre de grandes valeurs de ε , et la méthode est alors linéaire, tandis que pour $\delta = 0$ la méthode est équivalente à la méthode MVDR.

Les méthodes MVDR et WNC sont des améliorations de la méthode de Bartlett qui portent sur la façon dont les données sont comparées au replica. On peut également jouer sur la définition du replica. L'expansion géométrique prise en compte dans l'équation (5.2) via le terme en $1/r_i^s$ n'est pas indispensable, la méthode étant basée sur les accords de phase et non d'amplitude. Il est parfois plus efficace de ne considérer que l'exponentielle (*i.e.*, que le terme de phase) dans le replica. Suivant le même principe, on peut se limiter à la phase des enregistrements en normalisant la matrice K de façon à ne garder que sa phase.

Dans la pratique, ces méthodes sont appliquées à plusieurs fréquences, et les cartes de probabilité P obtenues pour chacune d'entre elles sont sommées de façon incohérente afin d'augmenter le rapport signal sur bruit du résultat, selon la racine carrée du nombre de fréquences utilisées.

Un autre approche serait de traiter de façon cohérente les cartes obtenues pour chaque fréquence, comme décrit par Debever and Kuperman [2007]. Un pré-traitement serait alors nécessaire, qui consisterait à les corrélérer par un capteur de référence afin de créer une source virtuelle à l'emplacement de ce capteur, comme décrit dans les chapitres précédents de ce manuscrit. Cette opération est nécessaire afin d'assurer que toutes les fréquences utilisées aient la même référence de phase. Le rapport signal sur bruit des cartes de probabilité serait alors meilleur, puisque le gain serait proportionnel au nombre de fréquences utilisées et non plus à la racine carrée de ce nombre. L'inconvénient de cette méthode est qu'il faut traiter toutes les fréquences en même temps, ce qui ajoute une dimension à la matrice K (dans la pratique cette matrice doit rester de dimension 2 pour le traitement, mais son nombre d'éléments est multiplié par le nombre de fréquences considérées, comme décrit par Debever and Kuperman [2007]). La matrice K devient alors trop grosse pour être traitée par les ordinateurs courants, dès lors que l'on considère un nombre de fréquences permettant une amélioration significative du rapport signal sur bruit. Une telle approche nécessite des ressources techniques particulières, ce qui limite son utilisation. Ne disposant pas des moyens adéquat, je n'ai pas utilisé les approches cohérentes.

► 5.1.2 Décomposition en valeurs singulières

La densité spectrale croisée K peut être décomposée selon ses M valeurs singulières λ_i . Si on note v_i les vecteurs singuliers associés, on a

$$K = \sum_{i=1}^M K_i = \sum_{i=1}^M \lambda_i v_i v_i^* \quad (5.8)$$

$$K v_i = \lambda_i v_i \quad . \quad (5.9)$$

On peut choisir de classer les valeurs singulières par ordre décroissant en valeur absolue. Dans un cas idéal, cette décomposition permet de séparer les composantes du champ d'onde résultant des différentes sources de bruit. Cette décomposition peut être vue comme une projection sur une base de fonctions de Green pour chaque point source possible (dans la pratique, la matrice K a un rang au mieux égal au nombre de capteurs, et la base ne peut décrire toutes les positions possibles tout

en étant orthogonale). Chaque vecteur singulier v_i est alors la composante de Fourier à la pulsation ω de la fonction de Green entre une source de bruit particulière et le récepteur i . Dans la pratique, il arrive que cette décomposition ne soit pas parfaite, et qu'un vecteur singulier corresponde à plusieurs sources, ou de la même façon qu'une source soit décrite par plusieurs vecteurs singuliers. Cela arrive principalement lorsque deux sources sont proches spatialement, ou lorsque l'une d'entre elles a une énergie très supérieure aux autres.

Cette décomposition est intéressante car elle permet d'appliquer les méthodes de formation de voie à chacune des matrices K_i issues de la décomposition de K . On peut alors retrouver la position de chacune des sources de bruit de façon indépendante. Cette méthode est particulièrement utile lorsqu'on est en présence d'une source de bruit dominante. On ne peut alors pas toujours étudier les sources secondaires car celles-ci sont masquées par la première. La décomposition en valeurs singulières permet de supprimer (ou au moins d'atténuer) la première source afin d'étudier ces sources secondaires.

► 5.1.3 A propos du rang de la matrice de densité spectrale croisée

Pour pouvoir utiliser les algorithmes de formation de voie non linéaires (MVDR ou WNC), il faut que la matrice de densité spectrale croisée $K(\omega)$ soit inversible afin de pouvoir former les replicas des équations (5.4) et (5.7). De même, pour pouvoir utiliser la décomposition en valeurs singulières, il faut que K soit de rang au moins égal au nombre de valeurs singulières que l'on veut séparer. Dans la pratique, si on calcule K à l'aide de l'équation (5.1), cette matrice sera de rang 1 et on ne pourra utiliser ni les algorithmes non linéaires ni la décomposition en valeurs singulières.

Il faut donc calculer K de façon à obtenir une matrice de rang plus élevé. J'ai utilisé pour cela deux techniques, donnant deux matrices K^a et K^b pouvant remplacer K dans le calcul de la réponse de la formation de voie. La première consiste à découper les enregistrements en N_s segments temporels de même durée, avec éventuellement un recouvrement des segments. En prenant la transformée de Fourier de chacun de ces segments, on obtient des vecteurs $d_i^k(\omega)$, i représentant le numéro du récepteur et $k \in 1 \dots N_s$ le numéro du segment. La matrice $K^a = (K_{ij}^a)$ peut alors être construite comme suit :

$$K_{ij}^a(\omega) = \sum_{k=1}^{N_s} d_i^k(\omega) d_j^{k*}(\omega) \quad . \quad (5.10)$$

Cette matrice est alors de rang N_s dès lors que l'on suppose les vecteurs d_i^k orthogonaux lorsqu'on fait varier k , ce qui est le cas en pratique.

Une autre façon de rendre K inversible, c'est-à-dire de s'assurer qu'aucune de ses valeurs propres n'est nulle, est de lui ajouter un bruit α sur la diagonale :

$$K_{ij}^b(\omega) = d_i(\omega) d_j^*(\omega) + \alpha I \quad , \quad (5.11)$$

où I est la matrice identité. Il faut alors choisir α petit de telle façon que le replica ne soit pas affecté par la bruit rajouté.

► 5.1.4 Ambiguïté du modèle de vitesse

Comme nous l'avons vu, les méthodes de formation de voies permettent de trouver la position de la (ou des) source de bruit pour un modèle de propagation donné. Si on change le modèle de propagation, la position de la sources trouvée à l'aide de ces méthodes change aussi. En effet, ces méthodes sont basées sur la concordance des phases entre les signaux mesurés et calculés, et font donc intervenir des rapport entre les distances source-récepteur et la vitesse. Une mauvaise description du milieu en terme de vitesses implique ainsi une mauvaise localisation de la source afin de compenser l'erreur introduite et de faire coïncider au mieux ces phases. Il faut donc connaître parfaitement le milieu (*i.e.*, la vitesse de propagation des ondes) afin de pouvoir déterminer la position réelle des sources, ce qui est rarement le cas en pratique.

Lorsque l'on ne connaît pas le modèle de vitesse dans le milieu, on peut toutefois l'inverser conjointement avec la position de la source. Pour cela, on calcule les cartes de probabilité de présence de la source pour différents modèles de vitesses. La carte ayant l'amplitude maximale la plus élevée (*i.e.*, donnant la probabilité de présence la plus forte) correspond au modèle de vitesse le plus proche du modèle réel. C'est l'approche utilisée pour la mesure des courbes de dispersions par les méthodes en fréquence/nombre d'onde présentées au paragraphe 3.1.2.

5.2 Étude de la vallée de Waimangu, Nouvelle-Zélande

► 5.2.1 Présentation de la vallée de Waimangu

La vallée de Waimangu est une région hydrothermale dans la zone volcanique de Taupo, sur l'île du Nord de la Nouvelle-Zélande. Cette vallée est située à l'extrémité sud du complexe volcanique Okataina, à 26 km au sud de Rotorua. Le système géothermal de Waimangu a été formé lors de l'éruption violente du Mont Tarawera le 10 juin 1886. Cette éruption fut relativement petite en terme de quantité de matériel éjecté (entre un et deux kilomètres cube de lave), tout en étant la plus grosse jamais vue en Nouvelle-Zélande depuis l'installation des européens au milieu du 17^e siècle [Keam, 1988]. L'énergie alors libérée fut telle que l'éruption engendra, en 6 heures seulement, 7 cratères le long des 18 kilomètres de la ligne de fracture. Après l'éruption, une activité hydrothermale de surface est apparue dans la zone de Waimangu. Quelques années plus tard, c'est dans cette zone que le plus grand geyser jamais connu fut actif, durant les années 1900 à 1904, pouvant atteindre 450 m de haut, avec un cycle d'une durée moyenne de 36 heures. L'activité actuelle est limitée à la zone des lacs Frying Pan et Inferno. Elle est constituée de fumerolles et de petits geysers, ainsi que d'une activité cyclique de ces deux lacs [Scott, 1994]. Le lac acide Inferno, d'une centaine de mètres de diamètre, voit ainsi son niveau varier en quelques heures de 10 m, avant de déborder, dans un cycle d'une période moyenne de 38 jours (voir la figure 5.1 et [Scott, 1992, 1994]) mettant en jeu 60 000 m³ d'eau. Les variations de la température du lac associées à ce cycle est de l'ordre de 30°C. Des études précédentes ont montré une signature sismique du cycle [Stanton, 1978, Tosha et al., 1996], comme c'est le cas au geyser Old Faithful en Californie [Kedar et al., 1996, 1998]. Elles montrent qu'un bruit acoustique est généré pendant la phase de montée du cycle, probablement à cause de la cavitation de bulles [Vandemeulebrouck et al., 2005]. L'étude du bruit sismique dans

de telles zones, et en particulier la localisation des sources de bruit, est donc une étape vers la compréhension des mécanismes mis en jeu.

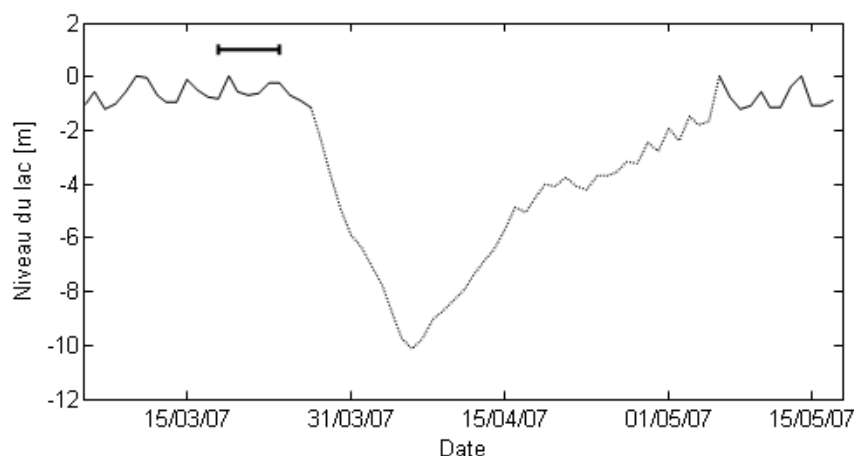


FIG. 5.1 – Niveau du lac Inferno pour la période du 5 mars au 16 avril 2007, représentant un cycle complet d'activité du lac Inferno. Le trait noir indique la période de déploiement (voir le tableau 5.1). Durant la période tracée en pointillés le fonctionnement du capteur de pression dans le lac était erratique, et les données utilisées pour la figure sont celles d'un autre cycle de même période et amplitude.

Afin d'étudier l'activité sismique de la zone de Waimangu, nous avons déployé des réseaux constitués de 16 géophones de fréquence de coupure 4,5 Hz. 3 zones de déploiement ont été retenues, comme le montre la figure 5.2a. Sur cette figure, la zone 2 correspond à l'emplacement où se trouvait le geyser de Waimangu (Old Geyser Site). La figure 5.2b montre la géométrie des réseaux utilisés. Les méthodes mises en œuvre ne nécessitent pas de géométrie particulière, et nous avons déployé les réseaux selon les contraintes dues au milieu (végétation, relief), en essayant d'obtenir les ouvertures de réseaux maximales autorisées par notre dispositif expérimental.

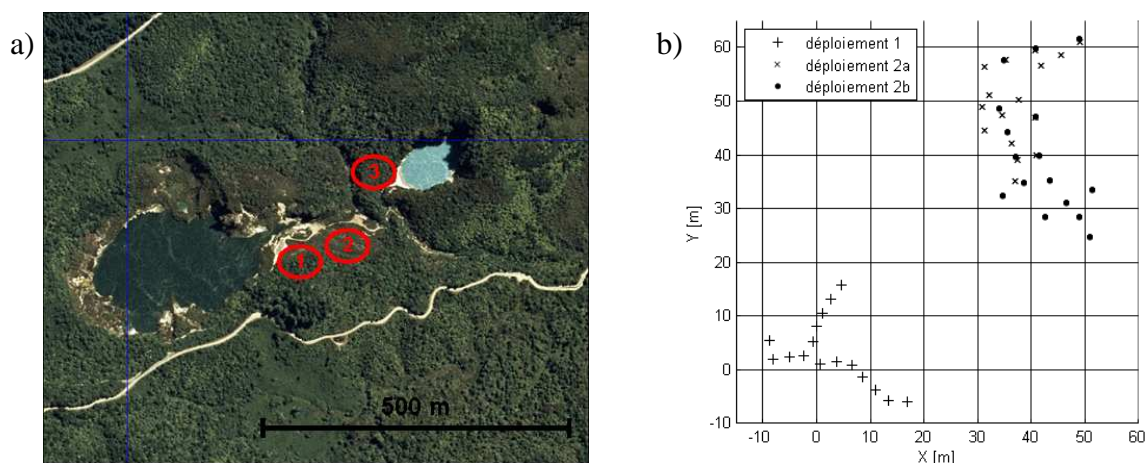


FIG. 5.2 – **a.** Image satellitaire de la zone de Waimangu. On voit sur la gauche le lac Frying Pan et sur la droite le lac Inferno. Les numéros indiquent les positions de déploiement des réseaux (deux réseaux ont été déployés à la position 2). **b.** Carte des 3 réseaux déployés aux emplacements 1 et 2 de la figure a. Les réseaux sont tous constitués de 16 géophones.

Nous avons pu faire des acquisitions à différents moments du cycle d'Inferno, comme le montre

la table 5.1. Pendant toute la période des acquisitions, le lac Inferno était à un niveau haut de son cycle, comme indiqué sur la figure 5.1. Cependant, le niveau du lac oscille autour de ce niveau haut, pouvant provoquer son débordement. Pour chaque déploiement, différents enregistrements de bruit ambiant ont été effectués, de durées comprises entre 30 et 300 s.

Jour (mars 2007)	18	19	19	20	21	21	22	22	23	24
Zone de déploiement	1	1	2a	2a	2a	3	3	2b	2b	2b
Variation du niveau d'eau	↘	↗	↗	↘	↘	↘	↘	↘	↗	↗

TAB. 5.1 – Récapitulatif des différentes acquisitions. La position des zones de déploiement est indiquée sur la figure 5.2a. Le sens de variation du niveau du lac Inferno correspondant aux oscillations autour de sa position haute est indiqué pour chaque acquisition.

Je n'ai pas pu exploiter les données relatives à la troisième zone de déploiement à cause de problèmes techniques d'acquisition générant un bruit électronique important dans les enregistrements, ce qui rend impossible la localisation des sources de bruit sismique à cause d'une baisse du rapport signal sur bruit (*i.e.*, bruit cohérent sur bruit incohérent).

L'étude présentée ici a donné lieu à deux publications [Legaz et al., 2009, Vandemeulebrouck et al., 2009].

► 5.2.2 Influence du cycle

Stanton [1978] et Tosha et al. [1996] ont observé que le cycle des lacs avaient une signature sismique. J'ai vérifié cette observation en calculant, pour chaque acquisition de bruit sismique pour les deux premières zones de déploiement, l'énergie des signaux (RMS). Ces deux zones de déploiement étant proches et ayant le même type de sol, on peut supposer que le type est la répartition de sources de bruit autour des réseaux est comparable, et qu'on peut donc comparer les énergies enregistrées sur chacun des deux réseaux. Ces énergies sont tracées sur la figure 5.3, en fonction des dates d'acquisition. Sur cette figure sont également indiquées les variations (en échelle arbitraire adaptée aux variations d'énergie) de température et de niveau du lac Inferno, qui permettent de suivre le cycle. Les énergies sont calculées sur les signaux filtrés entre 5 et 15 Hz et moyennées sur les 16 géophones du réseau pour chaque enregistrement. Nous retrouvons ici le résultat des études précédentes montrant une signature sismique du cycle qui se traduit par une augmentation de l'énergie du bruit quand le niveau du lac Inferno monte.

La question qui se pose alors est de savoir si le changement d'énergie du bruit est lié à une augmentation de l'activité des sources de bruit présentes quand le niveau du lac baisse, ou si elle est due à l'apparition de nouvelles sources lors du changement de régime du système hydrothermal. Pour répondre à cette question, j'ai utilisé la méthode de formation de voie décrite précédemment afin de localiser les sources de bruit autour d'un réseau (correspondant au déploiement 2a, voir figure 5.2) pour deux périodes différentes correspondant à une montée ou non du lac. Ne connaissant pas le modèle de vitesse, nous avons utilisé un modèle homogène, dont nous avons fait varier la vitesse afin de trouver la meilleure probabilité de présence de la source. La vitesse ainsi retenue est de 260 m/s. La figure 5.4 montre les réponses de la formation de voie pour les deux périodes.

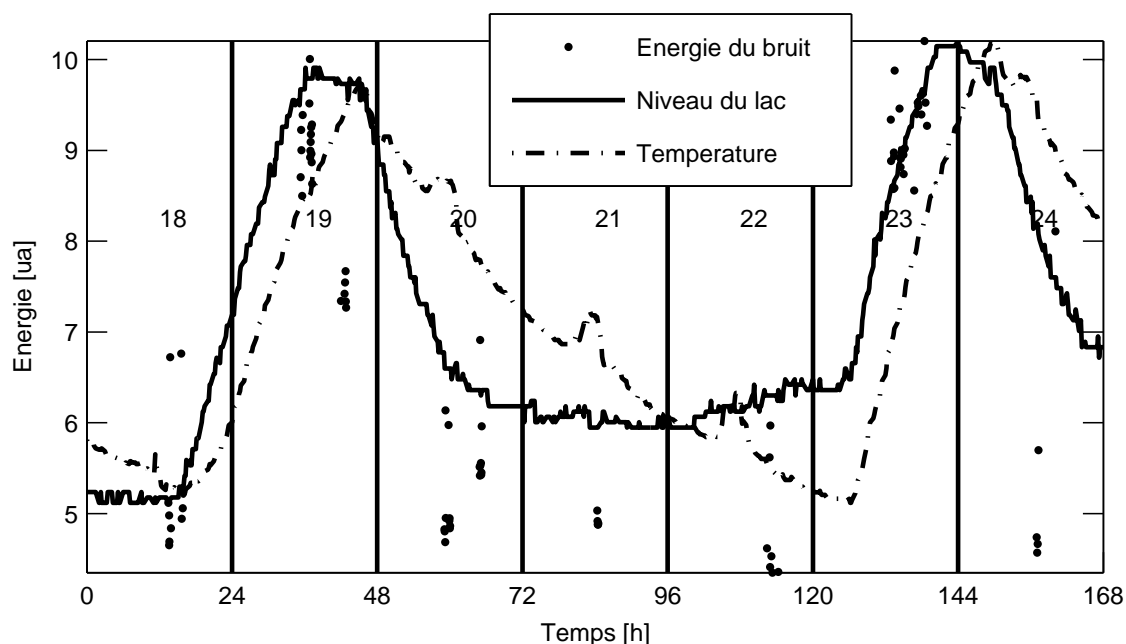


FIG. 5.3 – Énergie des signaux de bruit filtrés entre 5 et 15 Hz moyennée sur les 16 capteurs de chaque réseau. Afin de suivre les oscillations du niveau lac Inferno au cours du temps, les variations de niveau et de température dans le lac sont aussi indiquées, avec des d'amplitudes arbitraires ajustées aux variation d'énergie du bruit. Les numéros indiquent les jours du mois de mars 2007. Seuls les déploiements dans les zones 1 et 2 sont utilisés afin de comparer des situations proches (même type de sol, distances aux sources comparables). Chaque point représente une acquisition de 30 à 300 s de durée. La dispersion des résultats peut être expliquée par la présence d'événements « sismiques » tels que le passage de personnes à proximité du réseau.

Lorsque le niveau du lac baisse, une seule source de bruit semble dominer les enregistrements. Cette source est localisée proche de la surface. Quand le lac entre dans une phase de montée, la distribution des sources change et plusieurs sources apparaissent, ce qui rend la localisation plus difficile. On voit sur la figure 5.4b une source située plus en profondeur et ayant une extension spatiale plus importante que la source de la figure 5.4a. L'extension spatiale de cette source peut cependant être biaisée, l'ouverture du réseau étant trop faible pour avoir une résolution suffisante à cette profondeur. La source présente lorsque le niveau du lac baisse est toujours visible après le changement de régime, même si elle n'est plus dominante. Afin de vérifier que c'est bien la même source que l'on retrouve quand le niveau monte, j'ai appliqué la technique de séparation de sources par décomposition en valeurs singulières (SVD). La figure 5.5 montre la réponse de la formation de voie appliquée aux deux premiers vecteurs singuliers issus de la SVD de la matrice de densité spectrale croisée K pour la période où le niveau du lac monte. Les deux sources de bruit ont bien été séparées, chacune correspondant à un seul vecteur singulier. La source la plus faible en énergie, difficilement visible sur la figure 5.4b, peut maintenant être étudiée plus facilement. On remarque en particulier que sa position correspond bien à celle de la source présente dans la période où le niveau du lac baisse. L'extension spatiale de cette source trouvée ici est plus importante que dans la période où le lac monte. Cette augmentation de largeur peut être attribuée à deux causes :

1. une imprécision dans la localisation de la source due à la faible énergie de cette source com-

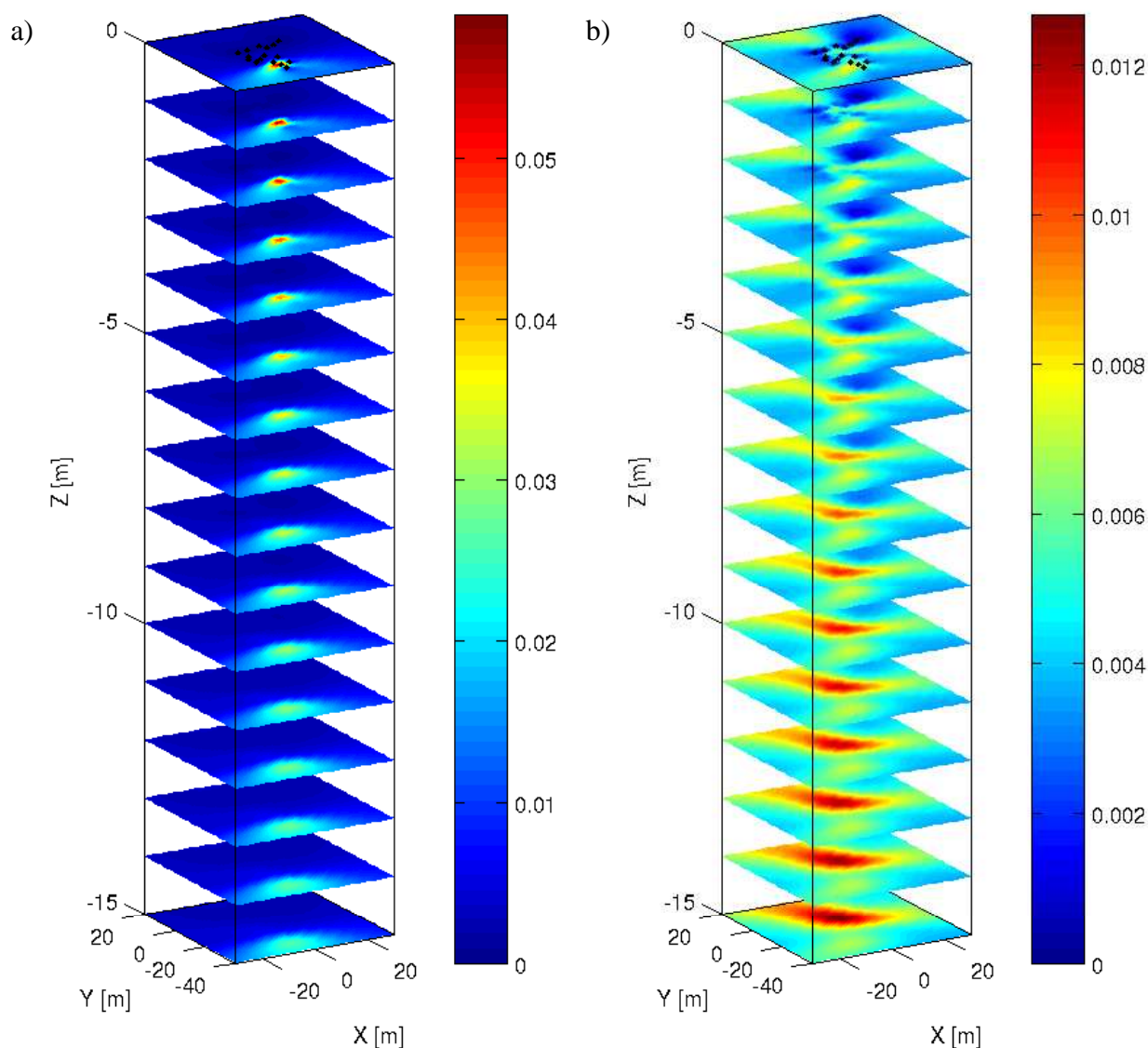


FIG. 5.4 – Réponse de la formation de voie (WNC, avec $\delta = -10$ dB) obtenue pour le réseau 2a, dans une période **a.** où le niveau du lac Inferno baisse ; et **b.** où le niveau du lac monte. Le modèle de vitesse utilisé est homogène et fixé à 260 m/s et les résultats sont calculés en utilisant les fréquences de 5 à 25 Hz par pas de 0,5 Hz. Les points noirs indiquent la positions des géophones.

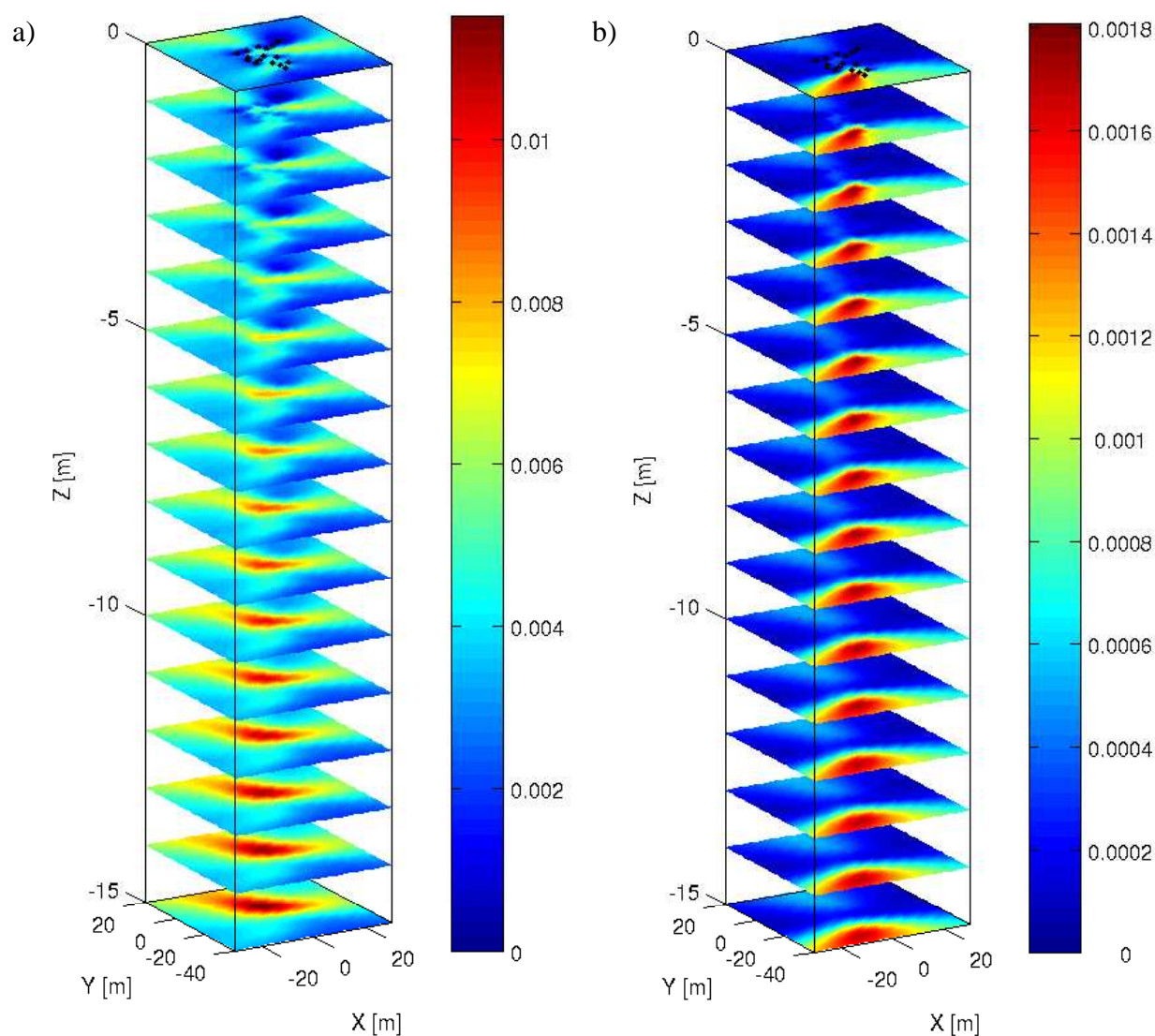


FIG. 5.5 – Réponse de la formation de voie appliquée à chacun des deux premiers vecteurs singuliers issus de la SVD de la matrice de densité spectrale croisée du champ enregistré lorsque le niveau du lac Inferno monte. Les deux sources visibles sur la figure 5.4b peuvent maintenant être étudiées séparément.

parée à la source dominante. La décomposition en valeurs singulières permet de séparer les sources, mais le rapport signal sur bruit de la source de faible énergie dans le second vecteur propre reste petit, ce qui nuit à une localisation précise de cette source.

2. un changement dans le mécanisme de la source lié au changement de régime du système géothermal.

La correspondance entre les sources des figures 5.4a et 5.5b n'est cependant pas remise en cause par cette différence d'extension spatiale. Cette information semble indiquer que le changement de régime du système géothermal associé à la montée du niveau du lac fait apparaître de nouvelles sources de bruit sismique, *a priori* liées à l'apparition de zones de cavitation et/ou un regain d'activité des geysers et des fumerolles environnantes suite à la montée de fluide, mais en maintenant l'activité des sources de bruit présentes avant cette phase. L'étude réalisée ici ne permet pas de conclure à propos d'une éventuelle modification de l'activité des sources de bruit présentes tout au long du cycle d'Inferno.

► 5.2.3 Traitement fréquentiel

Afin de mieux contraindre le modèle de vitesse et la position de la source de la figure 5.4a, nous avons réalisé une étude à plus haute résolution de la position des sources de bruit. L'algorithme utilisé est toujours le même, mais il est appliqué indépendamment à chaque fréquence, et pour une grille de recherche en espace et en vitesse beaucoup plus fine. Un premier calcul est effectué à une fréquence donnée (la fréquence choisie ici est 15 Hz), puis les bornes de recherches pour les fréquences voisines (14,9 et 15,1 Hz) sont définies en fonction de la position de source trouvée pour cette première fréquence. Le calcul est ainsi étendu vers les basses et les hautes fréquences, en définissant une grille de recherche centrée sur la position de la source et la vitesse du modèle trouvées pour la fréquence précédente (voir la figure 5.6). Cette façon de procéder permet de garder une grille très fine tout en limitant le domaine de recherche afin de minimiser le temps de calcul nécessaire. Les positions de source sont ainsi calculées indépendamment pour chaque fréquence (mis à part la définition de l'espace de recherche). L'optimum de vitesse en fonction de la fréquence constitue une mesure de la courbe de dispersion. Si on fait l'hypothèse que la source est la même pour toutes les fréquences, on obtient un critère pour définir le domaine de validité de la courbe de dispersion mesurée, soit de 10 à 28 Hz. Au delà de 28 Hz, la mesure de vitesse n'est plus fiable : deux maximums apparaissent dans le calcul de P , pour deux modèles de vitesses différents. La méthode de recherche basée sur la recherche d'un maximum n'est alors plus adaptée puisqu'elle peut suivre un maximum local. La raison du dédoublement du maximum de probabilité est difficile à déterminer précisément. Celui-ci peut provenir de l'apparition d'un mode harmonique de vitesse proche de celle du mode fondamental pour ces fréquences. Afin de vérifier cette hypothèse, il faudrait étendre le calcul de P à plus des modèles de vitesses afin de vérifier l'existence d'un tel mode.

► 5.2.4 Comparaison avec la polarisation spontanée

En parallèle des acquisitions de bruit sismique, des mesures de polarisation spontanée ont été effectuées au voisinage de la source identifiée sur la figure 5.4a, lorsque le lac Inferno est en train de

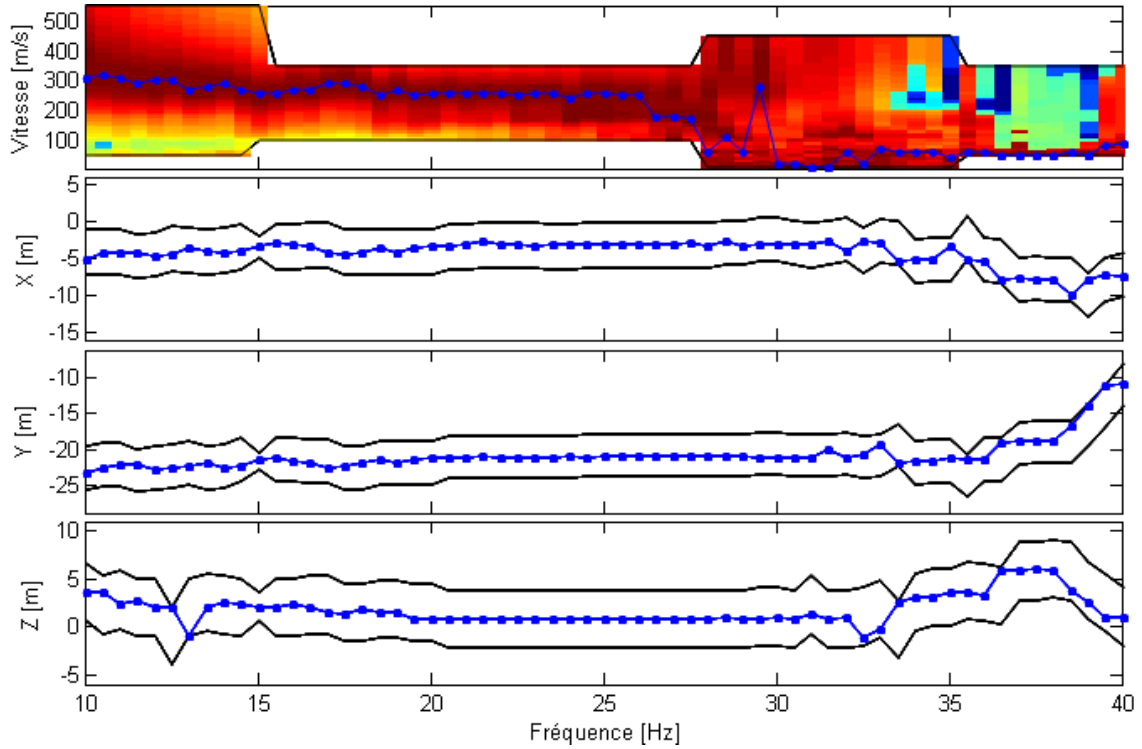


FIG. 5.6 – Modèle de vitesse et position de la source trouvés pour chaque fréquence. La figure du haut indique en couleur l'évolution du maximum de la réponse de la formation de voie en fonction du modèle de vitesse pour chaque fréquence. Les points bleus indiquent la position du maximum de probabilité $v_{opt}(\omega) = \max_v P(v, \vec{r}, \omega)$. Sur les autres figures, les points bleus indiquent les coordonnées spatiales du maximum de la probabilité, pour le modèle de vitesse donnant la plus forte probabilité de présence : $\vec{r}_{opt}(\omega) = \max_{\vec{r}} P(v_{opt}, \vec{r}, \omega)$. Les traits noirs de chaque figure indiquent les bornes de la grille de recherche.

baisser. Ces mesures ont été réalisées à l'aide d'électrodes *Petiau* Pb/PbCl₂ et d'un voltmètre à haute impédance (100 MΩ). Les mesures de polarisation spontanée fluctuent dans un intervalle 20 mV, sauf à la position de la source acoustique, où un pic étroit de 130 mV d'amplitude est observé (voir figure 5.7).

Les sources de polarisation spontanée peuvent être localisées selon un principe semblable à celui de la formation de voie décrite dans ce manuscrit. L'étude est effectuée sur la surface verticale définie par la position du profil de mesure et la profondeur. Cette surface, considérée homogène en terme de résistivité, est découpée en cellules contenant chacune un dipôle électrique vertical et un horizontal. Chacun de ces dipôles est alors excité et le champ électrique résultant est comparé aux anomalies de polarisation spontanée à la surface divisée par sa puissance [Patella, 1997, Revil et al., 2001].

Le champ électrique $E(\vec{r}, \vec{r}_s)$ engendré par un dipôle situé à la position \vec{r}_s et enregistré à la position \vec{r} s'écrit

$$E(\vec{r}, \vec{r}_s) = d \nabla G(\vec{r} - \vec{r}_s) \quad , \quad (5.12)$$

où G et d sont respectivement la fonction de Green (électrique) associée à l'équation de Poisson et le moment du dipôle. La densité d'intercorrélation P entre le champ $E_m(\vec{r})$ mesuré en \vec{r} et le replica

FIG. 5.7 – Position du réseau de géophones et mesures de polarisation spontanée. Le point de référence utilisé pour la polarisation spontanée est indiqué, ainsi que la position de la source acoustique localisée grâce à la figure 5.4a. L'échelle de couleur se réfère aux mesures de polarisation spontanée, en mV.

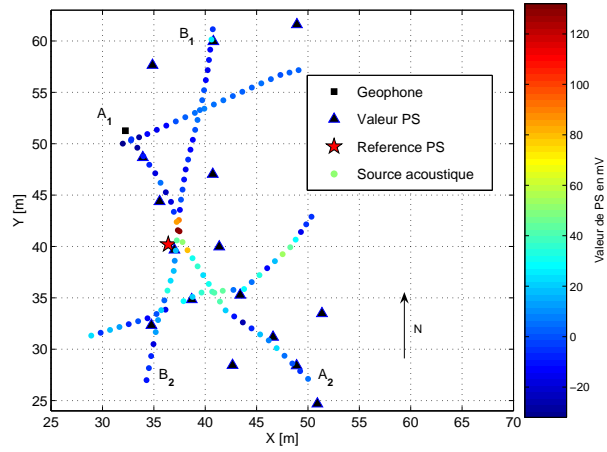
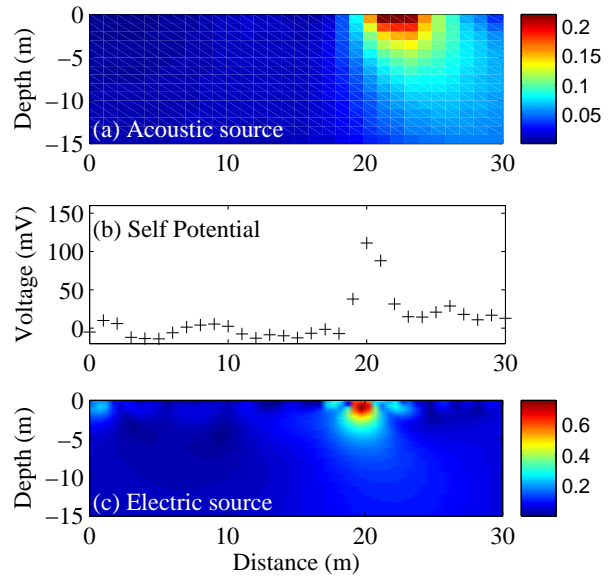


FIG. 5.8 – **a.** Résultat de la formation de voie extrait de la figure 5.4a pour la surface définie par l'axe vertical et le tracé du profil B de la figure 5.7. **b.** Mesures de polarisation spontanée le long de ce profil. **c.** Résultat de l'inversion pour la localisation du dipôle électrique. Les sources acoustique et électrique sont localisées en des endroits très proches.



théorique est alors une fonction vectorielle dont les composantes sont données par

$$P_x(\vec{r}_s) = C_x \int E_m(\vec{r}) \frac{\partial G(\vec{r} - \vec{r}_s)}{\partial x} dS \quad (5.13)$$

$$P_z(\vec{r}_s) = C_z \int E_m(\vec{r}) \frac{\partial G(\vec{r} - \vec{r}_s)}{\partial z} dS, \quad (5.14)$$

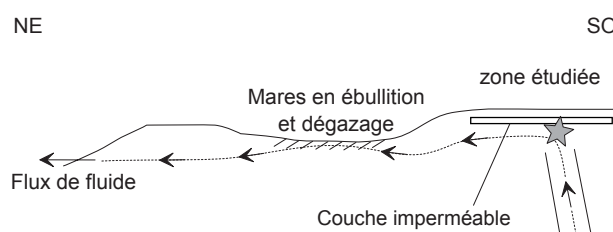
où C_x et C_z sont des constantes de normalisation prenant en compte la puissance du champ électrique [Iuliano et al., 2002], x étant l'abscisse le long du profil de mesure et z la profondeur. Le profil de mesure de polarisation spontanée considéré pour l'inversion est le profil B de la figure 5.7. La figure 5.8 montre la position du dipôle estimé par inversion.

Les positions trouvées pour les sources acoustique et électrique sont très proches, ce qui indique que ces deux sources sont issues de phénomènes prenant place dans la même structure localisée. Rien de visible en surface ne laisse supposer la présence de ces sources. A une quinzaine de mètres au nord se trouve une zone de fumerolles et de mares en ébullition (proche du géophone de coordonnées $(x = 37, y = 57)$ sur la figure 5.7), que l'on ne retrouve ni dans les données sismiques ni dans les données électriques. La source est localisée à une profondeur très faible, mais la précision sur cette profondeur est mauvaise à cause d'erreurs sur l'altitude des points de mesure, autant pour les

géophones que pour la polarisation spontanée, la végétation étant dense dans cette zone. La zone a donc été considérée comme plane, alors que la partie sud de la zone est située environ un mètre plus bas que la partie nord.

Une interprétation possible de cette double source acoustique/électrique est présentée sur la figure 5.9. Dans ce modèle très simple, la source électrique correspond à une remontée de fluide en

FIG. 5.9 – Modèle expliquant la présence simultanée d’une source acoustique et d’un dipôle électrique (indiqués par l’étoile grise), avec une fuite du fluide et ébullition et un dégazage en surface déportés par rapport à la position de ces sources.



ébullition qui, à la rencontre une couche imperméable, s’épand horizontalement et s’échappe dans une zone déportée latéralement. La rencontre du fluide sous pression avec la couche imperméable peut générer des ondes acoustiques [Shuli and Zhuo, 2003]. La couche imperméable peut être expliquée par la présence de dépôts de silice dûs à l’activité passée du geyser et/ou aux éruptions du cratère Echo situé à proximité de la zone. Des mesures d’émission de CO_2 montrent que le flux de gaz au voisinage de la position des sources est faible comparé aux fortes valeurs mesurées au niveau des mares en ébullition, ce qui est en faveur de l’hypothèse d’une couche imperméable. Des structures similaires comportant un flux vertical de fluide produisant de fortes anomalies de polarisation spontanée à la surface ont été identifiées sur le Mont Fuji [Aizawa et al., 2005], et sur d’autres volcans tels que le Mont Saint Helens [Bedrosian et al., 2008] ou encore sur l’île Vulcano en Sicile [Revil et al., 2008].

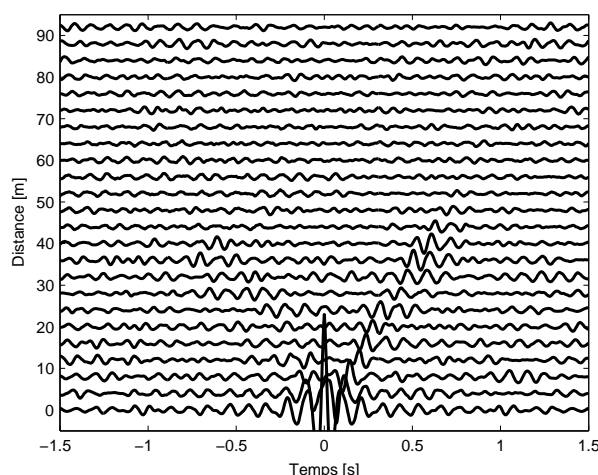
Cette étude est originale dans l’utilisation conjointe de données sismiques et de polarisation spontanée donnant des positions de sources acoustique et électrique très proches, et pouvant être interprétées comme issues du même système physique. Les mesures effectuées pour cette étude peuvent être réalisées en très peu de temps, et se révèlent être un outil intéressant pour l’analyse de systèmes à petite échelle. Cette approche pourrait mener à l’inversion jointe de données électriques et acoustiques, permettant une meilleure détermination de la géométrie des structures et des flux. Un autre axe d’étude serait le suivi temporel de ce site, afin de mieux caractériser les changements du système hydrothermal au cours du cycle.

5.3 Étude passive de la Solfatare, Italie

Lors de l’étude de la Solfatare, présentée au chapitre 4, nous avons réalisé des acquisitions de bruit pour chacun des profils déployés (voir le paragraphe 4.3). J’ai utilisé ces enregistrements continus pour rechercher la position des sources de bruit dans ce milieu fortement bruité.

Afin de tester la cohérence du bruit, j’ai corrélé, pour un profil particulier, les enregistrements en chacun des géophones par le signal enregistré sur le premier d’entre eux. On obtient ainsi une section de corrélation présentée sur la figure 5.10. Cette figure montre que le bruit enregistré n’est

FIG. 5.10 – Corrélation des enregistrements de bruit sur le second profil par le sismogramme du premier géophone de ce profil. Cette section montre que l'information cohérente ne peut être extraite au delà d'une cinquantaine de mètres de propagation.



pas cohérent sur toute la longueur du profil, c'est-à-dire que les signaux enregistrés en des points séparés de plus de 50 m ne sont pas issus des mêmes sources de bruit. La faible cohérence du bruit est due 1) à la forte atténuation dans le milieu ; 2) aux hétérogénéités qui peuvent masquer les signaux issus d'une source de bruit pour une position de récepteur donnée ; 3) à la multitude de sources présentes dans le milieu, chacune dominant les autres dans son entourage. La cohérence des signaux pourrait être améliorée en augmentant la durée d'enregistrement, mais ce n'est pas notre objectif ici. Le bruit ayant une distance de corrélation inférieure à la longueur des profils, et dépendant de la fréquence et du profil considérés, la formation de voie n'est pas efficace lorsque l'on considère l'ensemble des géophones du profil. En effet, les géophones étant proches les uns des autres enregistrent un bruit cohérent permettant une localisation de sources, tandis que ceux étant plus éloignés ajoutent du bruit dans la réponse de la formation de voie à cause de leur manque de cohérence. Afin d'améliorer le résultat, le réseau est découpé en sous réseaux, définis en déterminant une distance de cohérence pour le bruit enregistré localement, et la formation de voie est appliquée séparément pour chaque sous réseau. Les différentes images du milieu sont alors combinées pour donner la réponse du profil complet. La distance de cohérence peut être adaptée pour chaque partie du profil dans le cas de milieux fortement hétérogènes, comme c'est le cas ici. La procédure appliquée à chaque profil est la suivante :

- pour chaque récepteur considéré tour à tour comme récepteur de référence, la fonction de corrélation normalisée est calculée entre tous les récepteurs et le récepteur de référence ;
- la cohérence entre les traces est alors définie comme le maximum de la fonction de corrélation normalisée. Un seuil de cohérence minimal est défini (fixé à 0,2 dans cette étude), et les récepteurs ayant une cohérence avec la trace de référence inférieure à ce seuil ne sont pas considérés. Pour chaque capteur pris comme référence, on obtient ainsi un sous profil pour lequel les traces sont toutes cohérentes entre elles ;
- la méthode de formation de voie est appliquée indépendamment à chaque sous profil ;
- les réponses correspondant aux différents sous profils sont combinées afin de former la réponse du profil complet.

Le déploiement de chacun des profils n'étant pas simultané, la formation de voie est utilisée pour chaque profil indépendamment des autres, puis les réponses sont sommées pour obtenir la réponse globale de l'ensemble des profils. Le résultat du traitement, effectué dans différentes bandes

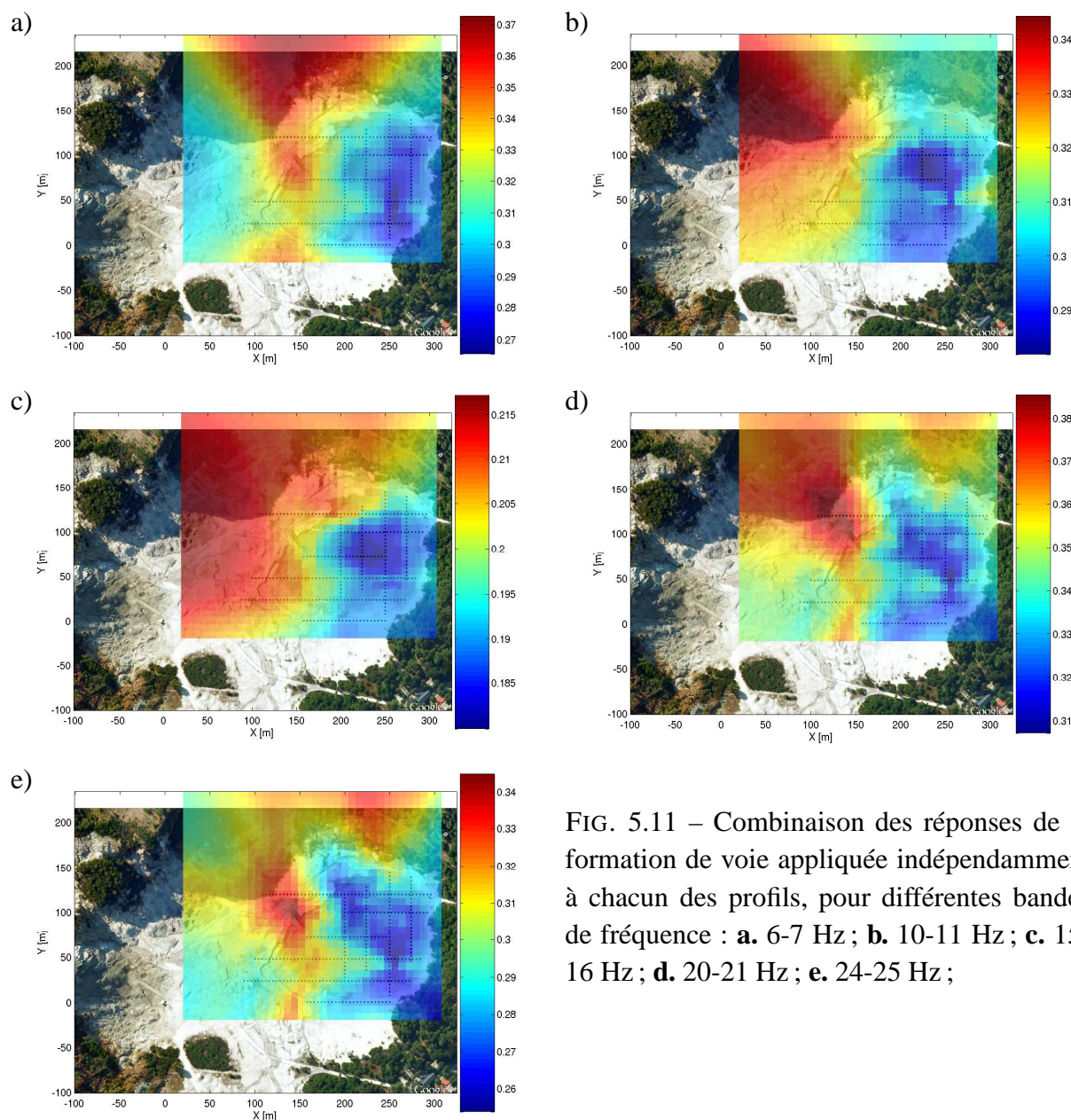


FIG. 5.11 – Combinaison des réponses de la formation de voie appliquée indépendamment à chacun des profils, pour différentes bandes de fréquence : **a.** 6-7 Hz ; **b.** 10-11 Hz ; **c.** 15-16 Hz ; **d.** 20-21 Hz ; **e.** 24-25 Hz ;

de fréquences, est présenté sur la figure 5.11. La position des sources de bruit varie avec la fréquence. Le bruit à basses fréquences provient principalement des fumerolles présentes à la périphérie du cratère, les plus importantes se trouvant dans le coin supérieur gauche des images présentées. Il est étonnant de constater que le mares en ébullition ne semblent pas produire de bruit sismique. La localisation des sources obtenue correspond à celle déjà trouvée par Bruno et al. [2007] et présentée sur la figure 5.12 (noter la différence d'orientation des figures, le nord étant vers le haut pour la figure 5.12, tandis que sa direction sur la figure 5.11 est indiquée sur la figure 4.16). Ces auteurs avaient caractérisé le bruit en mesurant son énergie en différents points du cratère, représentés par les triangles. Lorsque l'on considère les hautes fréquences, les sources obtenues sont plus locales. Leur positions correspondent aux zones chaudes et sèches de la figure 4.17. Ces zones ont un aspect croûté, avec des cavités situées à quelques centimètres sous le sol. Ces cavités pourraient jouer le rôle de caisses de résonance, ce qui expliquerait la localisation des sources de bruit à ces endroits.

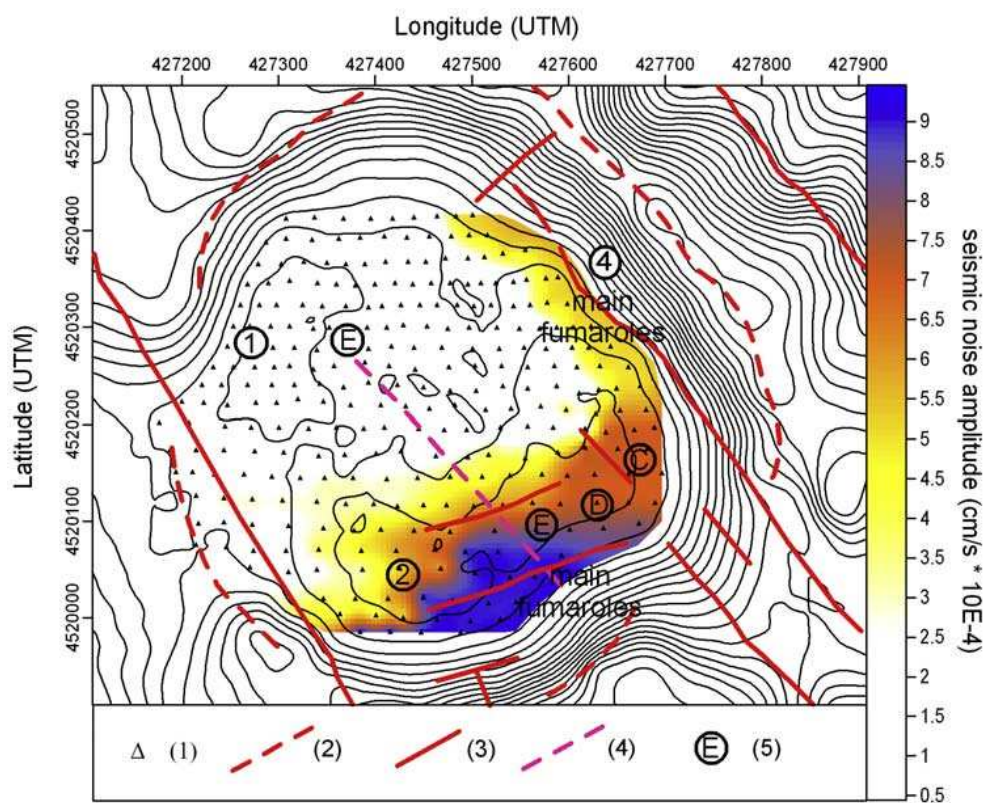


FIG. 5.12 – Distribution spatiale des amplitudes du bruit sismique ($\times 10^{-4}$ cm/s) dans la bande de fréquence 10-15 Hz dans le cratère de la Solfatare déterminée par Bruno et al. [2007] (figure extraite de l'article). La localisation de sources de bruit est la même que celle trouvée dans la bande 15-16 Hz par la formation de voie (figure 5.11 ; noter la différence d'orientation entre ces figures). Symboles : (1) grille des points de mesure ; (2) limites de la caldera ; (3) failles régionales.

La localisation des sources obtenue ici n'est pas aussi précise que celle obtenue dans l'étude de la vallée de Waimangu présentée plus haut. Cette différence de résolution s'explique par la proximité des sources étudiées dans la vallée de Waimangu, comparée à celles trouvées ici. Les sources étant plus lointaines, la courbure des fronts d'onde enregistrés sur le réseau est faible au vu de l'ouverture du réseau utilisé, et la distance source-réseau est ainsi difficile à retrouver. Une deuxième explication de la perte de résolution est la multitude de sources locales, combinée à la forte atténuation dans le milieu, qui diminue fortement la cohérence des signaux et empêche l'utilisation de grands réseaux. Un nouveau déploiement destiné à l'étude passive de cette zone est prévu en novembre 2008. Reprenant les conclusions de l'étude présentée ici, cette nouvelle expérimentation sera réalisée avec un matériel différent permettant une acquisition à plus basse fréquence, où l'atténuation est moins forte. Le réseau utilisé comportera beaucoup plus de points de mesure, ce qui permettra une acquisition simultanée des signaux de bruit dans l'ensemble du cratère.

Bibliographie chapitre 5

- K. Aizawa, R. Yoshimura, N. Oshiman, K. Yamazaki, T. Uto, Y. Ogawa, S. B. Tank, W. Kanda, S. Sakanaka, Y. Furukawa, T. Hashimoto, M. Uyeshima, and A. W. Hurst. Hydrothermal system beneath Mt. Fuji volcano inferred from magnetotellurics and electric self-potential. *Earth & Planetary Science Letters*, 235(1-2):343–355, Jun 2005. doi: 10.1016/j.epsl.2005.03.023.
- A. B. Baggeroer, W. A. Kuperman, and H. Schmidt. Matched field processing: source localization in correlated noise as an optimum parameter estimation problem. *The Journal of the Acoustical Society of America*, 83(2):571–587, 1988.
- A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky. An overview of matched fields methods in ocean acoustics. *IEEE Journal of Ocean Engineering*, 18(4):401–424, 1993.
- P. A. Bedrosian, M. J. Unsworth, and M. J. Johnston. Hydrothermal circulation at Mount St. Helens determined by self-potential measurements. *Journal of Volcanology and Geothermal Research*, 160:137–146, 2008.
- P. P. G. Bruno, G. P. Ricciardi, Z. Petrillo, V. Di Fiore, A. Troiano, and G. Chiodini. Geophysical and hydrogeological experiments from a shallow hydrothermal system at Solfatara Volcano, Campi Flegrei, Italy: response to caldera unrest. *Journal of Geophysical Research (Solid Earth)*, 112(B6):6201, Jun 2007. doi: 10.1029/2006JB004383.
- J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.
- H. Cox, R. M. Zeskind, and M. M. Owen. Robust adaptive beamforming. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35(10):1365–1376, Oct 1987.
- C. Debever and W. A. Kuperman. Robust matched-field processing using a coherent broadband white noise constraint processor. *The Journal of the Acoustical Society of America*, 122(4):1979–1986, Oct 2007. doi: 10.1121/1.2769830.
- T. Iuliano, P. Mauriello, and D. Patella. Looking inside Mount Vesuvius by potential fields integrated probability tomographies. *Journal of Volcanology and Geothermal Research*, 113(3):363–378, Mar 2002. doi: 10.1016/S0377-0273(01)00271-2.
- R. F. Keam. *Tarawera: the volcanic eruption of 10 June 1986*. R. F. Keam, 1988.
- S. Kedar, B. Sturtevant, and H. Kanamori. The origin of harmonic tremor at Old Faithful geyser. *Nature*, 379:708–711, Feb 1996. doi: 10.1038/379708a0.
- S. Kedar, H. Kanamori, and B. Sturtevant. Bubble collapse as the source of tremor at Old Faithful geyser. *Journal of Geophysical Research*, 103:24283–24300, Oct 1998. doi: 10.1029/98JB01824.
- W. A. Kuperman and G. Turek. Matched field acoustics. *Mechanical Systems and Signal Processing*, 11(1):141–148, 1997.

- A. Legaz, A. Revil, P. Roux, J. Vandemeulebrouck, P. Gouédard, T. Hurst, A. Bolève, A. Legaz, A. Revil, P. Roux, J. Vandemeulebrouck, P. Gouédard, T. Hurst, and A. Bolève. Self-potential and passive seismic monitoring of hydrothermal activity: A case study at Iodine Pool, Waimangu geothermal valley, New Zealand. *Journal of Volcanology and Geothermal Research*, 170(1-2): 11–18, Jan 2009. ISSN 0377-0273. doi: doi:10.1016/j.jvolgeores.2008.09.015.
- D. Patella. Self-potential global tomography including topographic effects. *Geophysical Prospecting*, 45:843–863, 1997.
- A. Revil, L. Ehouarne, and E. Thyreault. Tomography of self-potential anomalies of electrochemical nature. *Geophysical Research Letters*, 28(23):4363–4366, 2001.
- A. Revil, A. Finizola, S. Piscitelli, E. Rizzo, T. Ricci, A. Crespy, B. Angeletti, M. Balasco, S. Barde Cabusson, L. Bennati, A. Bolève, S. Byrdina, N. Carzaniga, F. Di Gangi, J. Morin, A. Perrone, M. Rossi, E. Roulleau, and B. Suski. Inner structure of La Fossa di Vulcano (Vulcano Island, southern Tyrrhenian Sea, Italy) revealed by high-resolution electric resistivity tomography coupled with self-potential, temperature, and CO₂ diffuse degassing measurements. *Journal of Geophysical Research*, 113:B07207, 2008. doi: 10.1029/2007JB005394.
- B. Scott. Characteristics of cyclic activity in Frying Pan and Inferno crater lakes, Waimangu. In *14th New Zealand Geothermal Workshop, Geothermal Institute, University of Auckland, New Zealand*, pages 253–258, 1992.
- B. Scott. Cyclic activity in the crater lakes of Waimangu hydrothermal system, New-Zealand. *Geothermics*, 23:555–572, 1994.
- W. Shuli and L. Zhuo. Water hammer phenomena in gas-water two-phase bubbly flow through a 90-degree bend tube. *Journal of Fluids Engineering*, 125:736–737, Jul 2003.
- T. Stanton. Seismic noise characteristics of Inferno Crater Lake, Waimangu. Msc thesis, Auckland University, Auckland, 1978.
- T. Tosha, T. Ishido, Y. Nishi, N. Matsushima, and B. Scott. Seismological and self-potential surveys at Inferno crater lake, Waimangu geothermal field. In *18th New Zealand Geothermal Workshop, Geothermal Institute, University of Auckland, New Zealand*, pages 249–254, 1996.
- J. Vandemeulebrouck, D. Stemmelen, T. Hurst, and J. Grangeon. Analogue modeling of instabilities in crater lake hydrothermal systems. *Journal of Geophysical Research (Solid Earth)*, 110 (B2):B02212, Feb 2005. doi: 10.1029/2003JB002794.
- J. Vandemeulebrouck, P. Gouédard, A. Legaz, P. Roux, A. Revil, A. W. Hurst, A. Jardani, and A. Bolève. Application of seismic noise and self-potential localization techniques to a buried hydrothermal vent (Waimangu Old Geyser site, New-Zealand). *Geophysical Journal International, Fast Track*, submitted, 2009.

Conclusion et perspectives

Mon travail a porté sur la méthode de reconstruction des fonctions de Green par corrélation, appliquée aux petites échelles de la géophysique (de la dizaine de mètres au kilomètre) et donc aux hautes fréquences (de 5 à 100 Hz) et à la sub-surface. L'objectif était d'investiguer les cas où cette reconstruction est imparfaite à cause d'une mise en défaut des hypothèses de base du théorème de corrélation, en particulier en ce qui concerne l'isotropie du champ et donc la distribution de sources de bruit ambiant. L'étude de l'influence de cette distribution a pu être réalisée grâce à un jeu de données réelles, ce qui n'avait jamais été effectué auparavant. J'ai montré qu'elle contrôle la convergence de la corrélation vers la fonction de Green en ce qui concerne les ondes directes. Dans le cas d'un milieu diffusant, l'utilisation de coda permet cependant de s'en affranchir totalement. Cette étude a également permis de montrer la faisabilité de la tomographie à partir de fonctions de Green reconstruites, ainsi que la possibilité de retrouver dans les fonctions de corrélation la décroissance de l'amplitude de la fonction de Green avec la distance de propagation. Une fois les cas de non convergence mis en évidence, mon travail a consisté à rechercher des méthodes permettant de retrouver malgré tout des informations sur le milieu de propagation, en prenant en compte les propriétés effectives du champ d'onde. J'ai ainsi montré qu'un bruit directionnel pouvait être utilisé pour reconstruire des fonctions de Green, moyennant une hypothèse de milieu homogène. Le traitement proposé dans ce manuscrit pour de tels cas ne peut cependant être appliqué que lorsque le bruit ne présente qu'une seule direction préférentielle. Le cas plus complexe d'un bruit multidirectionnel mais non isotrope n'a pas été abordé, alors qu'il est courant dans la pratique, en particulier aux échelles considérées dans cette thèse. Dans ce cas, une séparation des contributions de chaque source de bruit (comprendre groupe de sources provenant de la même direction) sera peut-être être indispensable, ce qui est envisageable via la décomposition en valeurs singulières par exemple.

La précision des mesures de dispersion obtenues à partir de fonctions de Green reconstruites a été étudiée. Pour cela, j'ai réalisé l'étude comparative d'un milieu connu via différentes méthodes. Les résultats obtenus montrent que les techniques de corrélation permettent d'obtenir, grâce à une interprétation différentes des données, des résultats parfois plus précis, ou à plus hautes fréquences que les méthodes classiques utilisant les mêmes données non corrélées. Les techniques de corrélation semblent donc prêtes pour une utilisation dans les études pratiques. Un autre aspect de ce travail a été de mettre en évidence les avantages à utiliser les techniques de corrélation, en particulier pour les applications liées à la sub-surface. L'abandon des sources synchronisées, et de l'équipement

d'acquisition nécessaire à leur mise en œuvre, en est un exemple. La possibilité de pouvoir étudier une zone entière à partir d'un nombre limité de points d'acquisition situés à sa périphérie, grâce à la reconstruction de tous les trajets entre ces points, est également un bénéfice important lié à l'utilisation de ces méthodes.

Bien que les techniques de corrélation soient aujourd'hui largement utilisées dans de nombreux domaines de la physique des ondes, leurs limitations ne sont pas toujours bien comprises. Il est encore difficile de prédire la convergence ou non de la fonction de corrélation vers la fonction de Green pour un jeu de données particulier. La définition même d'un critère de convergence est difficile. De nouvelles études théoriques et expérimentales sont ainsi nécessaires afin de mieux comprendre et anticiper les résultats. Un autre axe d'étude est le développement de traitements permettant l'accélération de cette convergence via par exemple l'utilisation de filtres auto-adaptatifs [Baig & Campillo, 2008]. Ce manuscrit s'est limité à l'étude des ondes de surface, les méthodes de corrélation étant énergétiquement plus favorables à leur reconstruction qu'à celle d'ondes de volume. L'étude du sous-sol par des méthodes basées sur les ondes de surface reste cependant à valider puisque, bien qu'elles aient montré de très bon résultats, il reste des cas où elles sont en désaccord avec les techniques plus classiques de réfraction ou de réflexion. La reconstruction d'ondes de volume par corrélation est aussi une importante et doit être investiguée. Il faudra alors être capable d'extraire ces ondes des fonctions de corrélation malgré leur rapport énergétique défavorable par rapport aux ondes de surface.

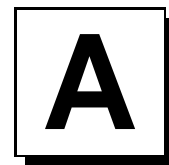
La dernière partie de ce manuscrit est consacrée à la caractérisation du bruit sismique ambiant, et en particulier à la localisation de ses sources à partir d'enregistrements continus. L'application des méthodes de formation de voie, classiquement utilisées en acoustique sous marine, a montré des résultats intéressants permettant une meilleure compréhension des mécanismes complexes mis en jeu dans les régions hydrothermales. Le matériel utilisé pour les études présentées dans ce manuscrit est léger et facile à déployer. Il est donc tout particulièrement adapté à l'étude de petites structures géologiques, mais ne permet pas l'analyse de phénomènes de plus grande envergure. Il a montré ses limites au cours de l'étude passive de la Solfatare. La poursuite de ce travail en utilisant un réseau ayant une ouverture plus grande et permettant des acquisitions à plus basses fréquences est déjà entamée, notamment avec le commencement d'une nouvelle thèse.

Le travail effectué au cours de ma thèse m'a amené à participer à des expérimentations regroupant simultanément plusieurs méthodes d'acquisition, portant sur des grandeurs physiques différentes, pour l'étude d'un même objet géologique. J'ai ainsi pu comparer des résultats de mesures sismiques et électriques, de polarisation spontanée, ou encore de température. Cette comparaison est très riche car elle permet une étude plus approfondie des mécanismes, chaque grandeur mesurée apportant une information différente sur les phénomènes physiques. Les analyses ont toujours été réalisées de façon indépendantes pour chaque grandeur physique, et la mise en commun des résultats n'a été qu'une dernière étape aidant à l'analyse du système. Une avancée majeure consisterait à pouvoir utiliser ces informations conjointement en une seule étape d'inversion. Le modèle alors obtenu serait beaucoup mieux contraint, et donc beaucoup mieux résolu, et la compréhension des phénomènes physiques plus aisée.

Le bruit ambiant pouvant être enregistré en continu, les méthodes sismiques passives peuvent être utilisées pour réaliser un suivi temporel de structures géologiques. Certains auteurs ont montré que la sensibilité aux variations du milieu est alors très importantes, ce qui a permis la mise en

évidence de changements de vitesse des ondes de l'ordre d'un pour mille, précurseurs d'éruptions volcaniques ou de séismes [Brenguier et al., 2008b,a, Duputel et al., 2008]. La résolution temporelle des variations dans le milieu est alors déterminée par le temps d'acquisition nécessaire à la convergence de la fonction de corrélation vers la fonction de Green. Ce temps de convergence dépend des échelles et de la complexité des structures, mais il est classiquement de quelques minutes pour des fréquences de 10 Hz et des distances de l'ordre de la dizaine de mètres, et de quelques jours à 0,1 Hz et des distances kilométriques.

Les algorithmes numériques mis en place au cours de cette thèse sont relativement simples. Cette remarque peut s'étendre à la quasi-totalité des traitements proposés par les différents auteurs utilisant les méthodes de corrélation. Ces méthodes peuvent ainsi être appliquées à de gros volumes de donnée avec des temps de calcul relativement courts. Les paramètres d'entrée sont de plus peu nombreux et essentiellement liés au pré-traitement des données, ce qui permet de ne pas avoir à les redéfinir pour chaque acquisition d'une campagne. Ces caractéristiques autorisent une utilisation automatique et systématique de ces méthodes, sans aucune intervention humaine. Cette constatation, mise en parallèle avec la possibilité d'un suivi temporel de structures géologiques, ouvre la perspective d'une utilisation des méthodes passives comme moyen d'alerte lié à des précurseurs d'éruptions volcaniques et de séismes. On peut également imaginer la création de bases de données de signaux corrélés, mis à la disposition de la communauté géophysique mondiale, comme le sont actuellement les enregistrements bruts continus. A plus petite échelle, un projet d'étude dynamique du modèle de vitesse d'un sol en fonction de l'infiltration de l'eau pluviale est en train de démarrer. Au vu de la fréquence des acquisitions nécessaires à cette étude, l'absence d'intervention humaine ni au moment de l'acquisition, ni au moment du traitement, est particulièrement intéressante.



Cross-correlation of random fields : mathematical approach and applications

Cette annexe est constituée de l'article [Gouédard et al., 2008d], dont un extrait a été présenté à la fin du chapitre 1. Cet article de revue propose dans une première partie une théorie mathématique la plus générale possible pour les méthodes de corrélation. Différentes applications sont ensuite décrites, à des échelles allant de quelques centimètres à plusieurs centaines de kilomètres.

Cross-correlation of random fields: mathematical approach and applications

P. Gouédard^{1*}, L. Stehly^{1,2}, F. Brenguier^{1,3}, M. Campillo¹, Y. Colin de Verdière⁴, E. Larose¹, L. Margerin⁵, P. Roux¹, F. J. Sánchez-Sesma⁶, N. M. Shapiro³ and R. L. Weaver⁷

¹Laboratoire de Géophysique Interne et de Tectonophysique, Observatoire de Grenoble, Université Joseph Fourier, Grenoble, France,

²CEA/DASE, Bruyères-le-Châtel, France, ³Laboratoire de Sismologie, Institut de Physique du Globe de Paris, CNRS, Paris, France, ⁴Institut Fourier, Université Joseph Fourier, Grenoble, France, ⁵CEREGE, Europôle Méditerranéen de l'Arbois, Aix-en-Provence, France, ⁶Instituto de Ingeniería, Universidad Nacional Autónoma de México, Coyoacán, Mexico, and ⁷Department of Physics, University of Illinois, Urbana, Illinois, USA

Received February 2007, revision accepted October 2007

ABSTRACT

Random field cross-correlation is a new promising technique for seismic exploration, as it bypasses shortcomings of usual active methods. Seismic noise can be considered as a reproducible, stationary in time, natural source. In the present paper we show why and how cross-correlation of noise records can be used for geophysical imaging. We discuss the theoretical conditions required to observe the emergence of the Green's functions between two receivers from the cross-correlation of noise records. We present examples of seismic imaging using reconstructed surface waves from regional to local scales. We also show an application using body waves extracted from records of a small-scale network. We then introduce a new way to achieve surface wave seismic experiments using cross-correlation of unsynchronized sources. At a laboratory scale, we demonstrate that body wave extraction may also be used to image buried scatterers. These works show the feasibility of passive imaging from noise cross-correlation at different scales.

INTRODUCTION

Traditional observational methods in seismology are based on earthquake records which results in two main shortcomings:

- 1 Most techniques are based on waves emitted by earthquakes that occurred only in geologically active areas, mainly plate boundaries. This results in a limited resolution in all other areas where earthquakes are not present. In particular, at stations far away from the source region, all the high-frequency information is lost due to the attenuation of the medium.
- 2 The occurrence of earthquakes is too low, preventing the study of real time change of active structures such as volcanoes or faults.

In the case of active seismic at smaller scales, the resolution is limited by the number and power of sources. It is thus difficult to image large areas or deep structures. Furthermore, controlled sources are difficult to carry out on hardly accessible places, like at the ocean bottom, where passive imaging could be much more convenient. For time-lapse monitoring, reproducible sources are necessary. This is very difficult to achieve for surveys of long duration, whereas noise wavefields may be stationary on these time scales.

Here we explore an alternative way of probing the Earth's interior using noise records only. The main idea is to consider seismic noise as a random source field when averaged over a long time series. In this particular case, cross-correlation between two stations yields the Green's function between these two points. As the seismic noise is mainly generated by atmospheric and oceanic forcing at the Earth's surface, the surface

*E-mail: pierre.gouedard@ujf-grenoble.fr

wave part of the Green's function is mostly extracted from the cross-correlation process.

At smaller scales, the same principle can be applied to study local structures. In this case, higher frequencies are used. At these frequencies, the wavefield is believed to be governed by local sources, which are unlikely to have the expected properties of randomness. This has to be taken into account in the processing.

In this paper, the theoretical relationship between noise cross-correlation and the Green's function is first discussed in section 2, based on theoretical derivations from (Colin de Verdière 2006a,b). Several applications are then presented, from large to small scales, using noise wavefields of different origin and physical properties: surface wave tomography at the regional scale in Western Europe (section 3) and at a more local scale in section 4 at the 'Piton de la Fournaise' volcano (Breguier *et al.* 2007); local P-waves extraction in the Parkfield network at the San Andreas Fault (Roux *et al.* 2005a) in section 5; site characterization using surface waves extracted from noise cross-correlation (Gouédard *et al.* 2006) in section 6; passive imaging of a buried scatterer at laboratory scale (Larose *et al.* 2006b) in section 7.

BACKGROUND AND MATHEMATICAL APPROACH

Historical background

The Green's function of a medium between two points A and B represents the record we would obtain at A if an impulsive source is applied at B.

In the case of a completely random wavefield, the cross-correlation of signals recorded between two points converges to the complete Green's function of the medium, including all reflection, scattering and propagation modes (Weaver 2005). To demonstrate this result and to define more precisely under which assumption it is valid, various experimental, numerical and theoretical approaches have been developed.

Historically speaking, helioseismology was the first field where ambient-noise cross-correlation performed from recordings of the Sun's surface random motion was used to retrieve time-distance information on the solar surface (Duvall *et al.* 1993; Gilles *et al.* 1997). The idea of day-light imaging was proposed by Claerbout (1968) in the context of prospecting. More recently, a seminal paper was published by Weaver and Lobkis (2001) that showed how diffuse thermal noise recorded and cross-correlated at two transducers fastened to one face of an aluminium sample provided the complete

Green's function between these two points. They theoretically interpreted this result by invoking equipartitioning of the modes excited in the aluminium sample. This result was generalized to the case where randomization is not produced by the distribution of sources, but is provided by multiple scattering that takes place in heterogeneous media (Lobkis and Weaver 2001).

The use of a spectral representation (Lobkis and Weaver 2001), the fluctuation-dissipation approach (Weaver and Lobkis 2001, 2003; van Tiggelen 2003; Godin 2007) or a correlation-type representation theorem (e.g. Wapenaar 2004) are rigorous theoretical approaches to interpret experimental results.

Experimental evidences demonstrated the feasibility of passive imaging in 1) acoustics (Lobkis and Weaver 2001; Weaver and Lobkis 2001; Larose *et al.* 2004), 2) seismology where Campillo and Paul (2003) retrieve the Green's function between two seismic stations from a collection of earthquakes, and 3) oceanography in shallow underwater acoustics where both direct and reflected wavefronts were retrieved from ambient-noise cross-correlation (Roux and Kuperman 2004; Sabra *et al.* 2005b). By summing the contributions of all sources to the correlation, it has been shown numerically that the correlation contains the causal and acausal Green's function of the medium (Wapenaar 2004). Cases of non-reciprocal (e.g. in the presence of a flow) or inelastic media have also been theoretically investigated (Wapenaar 2006; Godin 2007).

Derode *et al.* (2003a,b) proposed to interpret the Green's function reconstruction in terms of a time-reversal analogy and showed that correlation of multiply scattered waves could be used for passive imaging in acoustics. The convergence of the noise correlation function towards the Green's function in an unbounded medium can also be interpreted through the stationary phase theorem (Snieder 2004; Roux *et al.* 2005b).

In seismology, Aki (1957) proposed a long time ago to use seismic noise to retrieve the dispersion properties of surface waves in the subsoil. Shapiro and Campillo (2004) reconstructed the surface wave part of the Green's function by correlating seismic noise at stations separated by distances of hundreds to thousands of kilometres, and measured their dispersion curves at periods ranging from 5 to about 150 seconds. This method led to the first application of passive seismic imaging in California (Shapiro *et al.* 2005; Sabra *et al.* 2005a) with a much greater spatial accuracy than for usual active techniques. Larose *et al.* (2005) also used noise cross-correlation at small distances on the moon.

For the problem of elastic waves, it has been theoretically shown that the convergence of noise correlation to the

Green's function was bonded by the equipartition condition of the different components of the elastic field (Sánchez-Sesma *et al.* 2006a, 2007). In other words, the emergence of the Green's function is effective after a sufficient self-averaging process that is provided by random spatial distribution of the noise sources when considering long time series as well as scattering (Campillo 2006; Larose *et al.* 2006a).

The case of homogeneously distributed white noise sources

The scope of this section is to summarize the different theoretical approaches using mathematical tools that allow a global view of the correlation problem in any propagation medium. We will see that cross-correlation of noise recorded at two distant stations A and B yields the Green's function, assuming that the wavefield is a white noise distributed everywhere in the medium, with no assumption about the medium.

We consider any medium X, that does not need to be homogeneous, where the wave propagation equation is controlled by a damped equation that can be written as:

$$\frac{\partial^2 u}{\partial t^2} + 2a \frac{\partial u}{\partial t} - Lu = f \quad (1)$$

Here $a > 0$ is a constant that corresponds to the attenuation of the medium, $f(t, \vec{r})$ is the source field (i.e. the noise field in our case) and $u(t, \vec{r})$ denotes the displacement field. If $L = c^2(\vec{r}) \Delta$, we recognize the usual wave equation. In a more general calculation, L can be any negative self-adjoint elliptic differential operator. In more physical terms, L is an operator which preserves energy.

First of all we will introduce a definition of the Green's function in the frequency domain using the *integral kernel* of the operator L , and show that this definition is equivalent to the usual one. Then, by expressing the displacement field using the Green's function, we will calculate the cross-correlation and find how the derivative of the cross-correlation function is linked to the Green's function.

We introduce the integral kernel of an operator P , denoted by $\llbracket P \rrbracket(x, y)$ by:

$$\forall u : X \mapsto \mathbb{R}^3, (Pu)(x) = \int_X \llbracket P \rrbracket(x, y) u(y) dy$$

This is the 'continuous matrix' of the operator P . It has to be linked to the case of a finite space where one can define the matrix (P_{ij}) of P and write the following formula:

$$\forall u : X \mapsto \mathbb{R}^3, (Pu)_i = \sum_j P_{ij} u_j$$

We first consider a medium without attenuation, i.e. $a = 0$ in equation 1. Let us define the Green's function of L in the frequency domain, denoted by $\hat{G}(\omega + i\varepsilon, \vec{r}, \vec{r}_s)$, with ε a small positive value, as the opposite of the integral kernel of $((\omega + i\varepsilon)^2 + L)^{-1}$. In other words, \hat{G} is the *resolvent* of L evaluated at point $(\omega + i\varepsilon)^2$. The $\hat{\cdot}$ denotes a function defined in the Fourier space. ε ensures that $((\omega + i\varepsilon)^2 + L)$ is invertible as L has real eigenvalues. We will show that this mathematical definition of \hat{G} is the same as the usual one, which is the causal solution of the wave equation (equation 1) when the source function f is a Dirac impulse in time and space $\delta(t, \vec{r} - \vec{r}_s)$. The Green's function $\hat{G}(\omega + i\varepsilon, \vec{r}, \vec{r}_s)$ admits a limit as $\varepsilon \rightarrow 0^+$, denoted by $\hat{G}(\omega + i0, \vec{r}, \vec{r}_s)$, as a Schwartz distribution on the real axis. If L has a continuous spectrum, this limit is a smooth function (the 'limiting absorption principle'). \hat{G} can thus be written as:

$$\begin{aligned} \hat{G}(\omega + i0, \vec{r}, \vec{r}_s) &= -\llbracket ((\omega + i0)^2 + L)^{-1} \rrbracket(\vec{r}, \vec{r}_s) \\ &= -\int_X \llbracket ((\omega + i0)^2 + L)^{-1} \rrbracket \\ &\quad \times (\vec{r}, \vec{r}') \delta(\vec{r}' - \vec{r}_s) d\vec{r}' \\ &= -((\omega + i0)^2 + L)^{-1} \delta(\vec{r} - \vec{r}_s) \end{aligned}$$

which yields:

$$-((\omega + i0)^2 + L) \hat{G}(\omega + i0, \vec{r}, \vec{r}_s) = \delta(\vec{r} - \vec{r}_s)$$

The inverse Fourier transform of this equation gives a relation that is the usual definition of G in the case of a medium without attenuation:

$$\frac{\partial^2 G}{\partial t^2}(t, \vec{r}, \vec{r}_s) - L G(t, \vec{r}, \vec{r}_s) = \delta(t) \delta(\vec{r} - \vec{r}_s)$$

G is thus the solution of equation 1 in the case of an impulsive source in time and space. One can compute the inverse Fourier transform of $\hat{G}(\omega + i\varepsilon, x, y)$ using residue calculus, and take the limit as ε goes to 0 to obtain

$$G(t, \vec{r}, \vec{r}_s) = Y(t) \left[\frac{\sin t \sqrt{-L}}{\sqrt{-L}} \right] (\vec{r}, \vec{r}_s)$$

where Y is the Heaviside-step function, and where we denote $\sqrt{-L}$ the operator which eigenvalues are the images of the eigenvalues of L by the function $x \mapsto \sqrt{-x}$ (idem for the sinus function).

If we consider an attenuating medium, the Green's function $\hat{G}_a(\omega, \vec{r}, \vec{r}_s)$ is defined by the resolvent of L evaluated at point $\omega^2 + 2ia\omega$ instead of $(\omega + i0)^2$. It thus becomes

$$G_a(t, \vec{r}, \vec{r}_s) = Y(t) e^{-at} \left[\frac{\sin t \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}, \vec{r}_s) \quad (2)$$

We now define the time domain cross-correlation between the displacement at two points A and B as:

$$C(\tau, \vec{r}_A, \vec{r}_B) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T u(t, \vec{r}_A) \overline{u(t + \tau, \vec{r}_B)} dt \quad (3)$$

where the bar denotes the conjugate. $u(t, \vec{r})$ can be expressed using the Green's function G_a (here attenuation is necessary to ensure convergence of the integral, see Roux *et al.* (2005b)) and the source function f as follows:

$$u(t, \vec{r}) = \int_0^\infty dt' \int_X G_a(t', \vec{r}, \vec{r}_s) f(t - t', \vec{r}_s) d\vec{r}_s$$

We assume that f is a white noise distributed everywhere in the medium X , acting at any time t . In the frequency domain, a white noise contains all the frequencies with a random phase. In the time domain, this is a random wavefield such that the position and the activation time of each source are uncorrelated. In this case, and considering a damping medium, we replace the large T limit in the correlation by an ensemble average. We then obtain the following explicit expression for the correlation between the wavefields recorded at A and B (see Appendix A for mathematical details):

$$C(\tau, \vec{r}_A, \vec{r}_B) = \frac{\sigma^2 e^{-a|\tau|}}{4a} \left[(-L)^{-1} \left(\cos \tau \sqrt{-L - a^2} + a \frac{\sin |\tau| \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right) \right] (\vec{r}_A, \vec{r}_B) \quad (4)$$

where σ is the variance of the noise wavefield.

The time derivative of this equation is expressed in terms of the Green's function using (2), giving the more familiar expression:

$$\frac{d}{d\tau} C(\tau, \vec{r}_A, \vec{r}_B) = \frac{-\sigma^2}{4a} (G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)) \quad (5)$$

This means that for any medium, the time-derivative of the cross-correlation computed between the wavefields recorded at two stations A and B is the Green's function of the medium, provided that the damping coefficient is small enough and that noise sources behave as white noise acting everywhere in the medium. This is the same hypothesis as stated in Roux *et al.* (2005b), Lobkis and Weaver (2001) and others, but L is now an arbitrary negative definite elliptic operator, and so the present result is more general.

The case of a scattering medium

The previous calculation was made using sources randomly located anywhere and randomly active at any time. This is

a very strong hypothesis that is not valid in practical cases. Another demonstration of the link between cross-correlations and Green's functions can be made without any assumption about the noise sources location or their activation time. We only assume that there is equipartition at the boundaries of the region of interest, which means that each eigenmode is excited with the same level of energy.

A simple view of the relation between equipartition and correlation is given by the reconstruction of the Green's function of the homogeneous space using the azimuthal averaging of the correlation of plane waves, which are the eigenfunctions of the problem. Sánchez-Sesma and Campillo (2006) consider an isotropic distribution of P and S plane waves in an elastic medium. They found that the azimuthal average of the cross-correlation of motion between two points is proportional to the imaginary part of the exact Green's tensor between these points under the condition that the energy ratio S/P of the incident waves is the one predicted by equipartition. These results clearly show that equipartition is a necessary condition to retrieve the exact Green's function from correlations of the elastic field. In practice, one has to deal with complex media for which the eigenfunctions are unknown and therefore for which equipartition conditions cannot be explicitly specified in terms of local properties of the field.

Sánchez-Sesma *et al.* (2006a) discussed a particular case. They considered the field in the vicinity of a cylindrical scatterer embedded in an homogeneous space and illuminated isotropically with incident P and S plane waves in the ratio of equipartition of the homogeneous space. Taking into account the scattered waves, they showed that the azimuthal average of cross-correlations of motion between two points still yields the imaginary part of the exact Green's tensor of the heterogeneous medium, including the scattered waves, even at close distance from the scatterer. Is such a property still valid for any scattered or type of heterogeneity? What are the conditions required for the incident field? Weaver and Lobkis (2004) used an integral representation approach to study the problem of an heterogeneous region in an open medium. The essence of this property is expressed in the spectral theory of scattering that shows that the properties obtained in the simplest case of a homogeneous medium are formally valid in presence of heterogeneities. This is discussed in Colin de Verdière (2006a,b) as follows.

In the first step we will define the *spectral projector* and exhibit its expression using the cross-correlation function (equation 7). In a second step, will prove the relation between this projector and the Green's function, the so-called Stone formula (equation 8).

We denote $e_0(\vec{r}, \vec{k}) = e^{i\vec{k}\cdot\vec{r}}$ as the plane waves that are the eigenmodes of the homogeneous infinite space. In the case of a complex medium, the scattering theory (Ramm 1986; Reed and Simon 1978) tells us that, if the medium is heterogeneous only in a finite region, the eigenmodes in the whole space can be written as:

$$e(\vec{r}, \vec{k}) = e_0(\vec{r}, \vec{k}) + e_s(\vec{r}, \vec{k})$$

where e_s , the scattered waves, satisfies the so-called *Sommerfeld radiation* condition, which ensures that e_s will vanish when \vec{r} goes to infinity. This decomposition is still valid in the near field of the scatterers (i.e. inside the heterogeneous region).

For $I \subset \mathbb{R}_+$, we define the spectral projector of L on I , denoted by P_I , from its integral kernel $\llbracket P_I \rrbracket$ at any points \vec{r}_1 and \vec{r}_2 of X by:

$$\llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = (2\pi)^{-d} \int_{\lambda_k \in I} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^d \vec{k}| \quad (6)$$

where d is the dimension of the space and λ_k denotes the eigenvalue associated with the eigenfunction $e(\vec{r}, \vec{k})$. This is the projector on the sub-eigenspaces of L which eigenvalues are in I . Again, what P_I represents can be easily understood in the case of a finite number N of eigenvalues $\{\lambda_n\}$, where we can write:

$$\llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = \sum_{\substack{n \in [1, N] \\ \lambda_n \in I}} e_i(\vec{r}_1) \overline{e_n(\vec{r}_2)}$$

In this case, I represents a subset of values of $n \in [1, N]$ that are preserved, all the other being removed by the projector P_I . For example, if $u(\vec{r}) = \sum_{n=1}^N u_n e_n(\vec{r})$, we have:

$$(P_I u)(\vec{r}) = \sum_{\substack{n \in [1, N] \\ \lambda_n \in I}} u_n e_n(\vec{r})$$

We now will demonstrate that the derivative of the spectral projector on an interval around a value $\omega^2 \in \mathbb{R}_+$ is linked to the cross-correlation function at the corresponding pulsation ω . We thus consider an interval $I = [\omega_-^2, \omega_+^2]$ around ω^2 . The integral over $\lambda_k = c^2 |\vec{k}|^2 \in I$ in equation 6 defines a volume of integration of dimension d that can be decomposed into two integrals, one over a volume of dimension $d - 1$ defined by $c^2 |\vec{k}|^2 = \omega^2$ and the other over $|\vec{k}|$:

$$\llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = (2\pi)^{-d} \times \int_{c^2 |\vec{k}|^2 \in I} \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^{d-1} \sigma| |\vec{k}|^{d-1} d|\vec{k}|$$

where $|d^{d-1} \sigma|$ is the usual measure of the unit $(d - 1)$ -dimensional sphere. In the case $d = 3$, $d^2 \sigma$ is the infinitesimal solid angle. Taking the derivative with respect to ω_+ in this

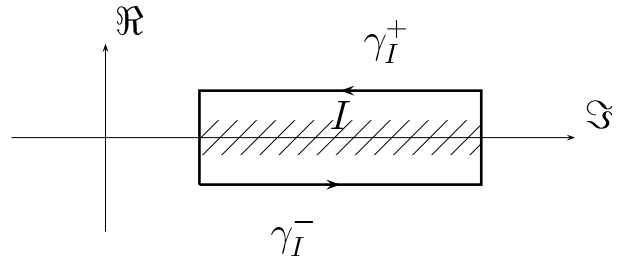


Figure 1 γ_I can be split into two contours γ_I^+ and γ_I^- which are complex conjugate and followed in opposite direction.

equation gives:

$$\frac{d}{d\omega_+} \llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = \frac{(2\pi)^{-d}}{c} \left(\frac{\omega}{c} \right)^{d-1} \times \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d^{d-1} \sigma|$$

In this formula, one can recognize the cross-correlation of random scattered waves of frequency ω recorded at points \vec{r}_1 and \vec{r}_2 that can be written as:

$$\hat{C}(\omega, \vec{r}_1, \vec{r}_2) = \frac{1}{\sigma_{d-1}} \int_{c^2 |\vec{k}|^2 = \omega^2} e(\vec{r}_1, \vec{k}) \overline{e(\vec{r}_2, \vec{k})} |d\sigma|$$

where σ_{d-1} denotes the total volume of the unit sphere in \mathbb{R}^{d-1} : $\sigma_0 = 2, \sigma_1 = 2\pi, \sigma_2 = 4\pi, \dots$

Using the two previous equations, we find

$$\frac{d}{d\omega} \llbracket P_I \rrbracket(\vec{r}_1, \vec{r}_2) = \frac{\sigma_{d-1}}{(2\pi)^d} \frac{1}{c} \left(\frac{\omega}{c} \right)^{d-1} C_\omega(\vec{r}_1, \vec{r}_2) \quad (7)$$

The projector P_I defined previously can also be written using the resolvent of the operator L using the Cauchy formula:

$$P_I = \frac{1}{2i\pi} \int_{\gamma_I} (L + \lambda)^{-1} d\lambda$$

where a γ_I is a contour in the complex plane which restriction to the real axis is I . This contour can be split into two contours defined by $\gamma_I^+ = \{\lambda \in \gamma_I \mid \Im(\lambda) \geq 0\}$ and $\gamma_I^- = \{\lambda \in \gamma_I \mid \Im(\lambda) < 0\}$ (\Im denotes the imaginary part) as seen in Fig. 1. As γ_I^+ and γ_I^- are followed in opposite directions and as they are complex conjugates, we obtain:

$$\begin{aligned} P_I &= \frac{1}{2i\pi} \int_{\gamma_I^+} [(L + \lambda)^{-1} - (L + \bar{\lambda})^{-1}] d\lambda \\ &= \frac{1}{\pi} \int_{\gamma_I^+} \Im(L + \lambda)^{-1} d\lambda \end{aligned}$$

and then, taking the integral kernel of this expression,

$$\begin{aligned}\|P_I\| &= \frac{1}{\pi} \int_{\mathcal{V}_e^+} \Im[(L + \lambda)^{-1}] d\lambda \\ &= -\frac{1}{\pi} \int_{\omega_-}^{\omega_+} \Im[(L + (\omega + i0)^2)^{-1}] 2\omega d\omega\end{aligned}$$

which leads to the Stone formula, using the definition of \hat{G}

$$\|P_I\|(\vec{r}_1, \vec{r}_2) = -\frac{2}{\pi} \int_{\omega_-}^{\omega_+} \omega \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) d\omega$$

This formula gives, by taking the derivative with respect to ω_+ ,

$$\frac{d}{d\omega_+} \|P_I\|(\vec{r}_1, \vec{r}_2) = -\frac{2\omega}{\pi} \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) \quad (8)$$

The combination of equation 7 and equation 8 finally gives:

$$\hat{G}(\omega, \vec{r}_1, \vec{r}_2) = -\frac{2^{d+1} \pi^{d-1}}{\sigma_{d-1}} \frac{c^d}{\omega^{d-2}} \Im \hat{G}(\omega, \vec{r}_1, \vec{r}_2) \quad (9)$$

This gives a generalization of equation 5 in the case of observation in a region without local sources, and requires no hypothesis about attenuation. It shows that the equipartition at boundaries of the region of interest is sufficient to obtain the Green's function from cross-correlation, whatever is the wave-field inside the medium is. This equation, established in the scalar case, can be extended to the elastic case using the same calculation. The cross-correlation function becomes a tensor, as well as the Green's function. Particular attention needs to be paid to velocities and dispersion relations as they depend on the type of waves.

The rate of convergence towards the Green's function

The question arises as to how much averaging is in principle necessary after which the Green's function is retrieved, and before which the cross-correlation remains dominated by noise. In the case that the field is due to uniformly distributed random sources, or in the case that the field is due to equipartitioned incident waves, it is possible to make theoretical estimates (Larose *et al.* 2004; Snieder 2004; Sabra *et al.* 2005c; Weaver and Lobkis 2005a). All these authors have concluded, not surprisingly, that the convergence proceeds like the square root of the amount of data used in the cross-correlation. signal-to-noise ratio, i.e. Green's function amplitude over residual fluctuations, is proportional to this square root. Quantitative estimates of the quality of the convergence are more challenging. Weaver and Lobkis (2005a) calculated the residual error in a scalar wave cross-correlation, and found it to be proportional to the energy in the diffuse field times the bandwidth times the fourth power of central frequency. A similar calculation

for closed systems was confirmed in laboratory measurements (Weaver and Lobkis 2005b). The residual error was compared to the amplitude of a ray arrival expected in the converged cross-correlation. Each ray arrival amplitude A depends on 1) the geometrical spreading of the Green's function, and 2) the spatial extension of the noise sources that coherently contribute to the Green's function reconstruction. This zone is characterized by a directivity angle $\delta\theta = \sqrt{\frac{c}{r\omega}}$, where r is the source-receiver distance. The ray arrival was shown to be apparent in the cross-correlation if $\delta t \delta\omega \gg A^{d-1}$, where δt is the amount of data record employed (this is often months in seismic applications), $\delta\omega$ is the bandwidth of interest (often around 1 Hz or less in seismic applications), and $A = r\omega/c$, (the source-receiver distance r times the wavenumber $k = 2\pi/\lambda$). The power is equal to one less than the dimension d of the propagation; thus $d - 1 = 1$ for Rayleigh waves. Propagation between distant source-receiver pairs, and propagation in three dimensions, are especially challenging to resolve, largely due to the weakness of such ray arrivals.

SURFACE WAVE TOMOGRAPHY OF EUROPE

Practically, cross-correlation can be used at different scales to image structures from noise. Here, we present an example of seismic noise processing to produce high-resolution Rayleigh and Love waves group velocity maps for a region surrounding the European Alps. We focused on the [5–50 s] period band, where surface waves are mostly sensitive to the crust.

Stehly *et al.* (2006) have shown that the seismic noise sources in the [5–20 s] period band cover a large surface when integrated over a long time. This allows us to retrieve the Green's function between two stations by correlating background seismic noise records. The emerging signal of the noise correlation function is dominated by surface waves, since the background seismic noise mainly consists of surface waves. The reconstructed Green's functions are stable over time and robust enough to measure surface wave propagation times with a precision of a few tenths of a second, independently of the azimuth of the considered station pair path (Stehly *et al.* 2007).

Passive imaging from seismic noise and Rayleigh wave group velocities was first used by Shapiro *et al.* (2005) and Sabra *et al.* (2005a) who provided images of the Californian crust. More recently, noise based surface-wave tomography has been applied in Tibet (Yao *et al.* 2006), New Zealand (Lin *et al.* 2007) and Korea (Kang and Shin 2006) and to produce large-scale Rayleigh wave group velocity maps across Europe (Yang *et al.* 2007).

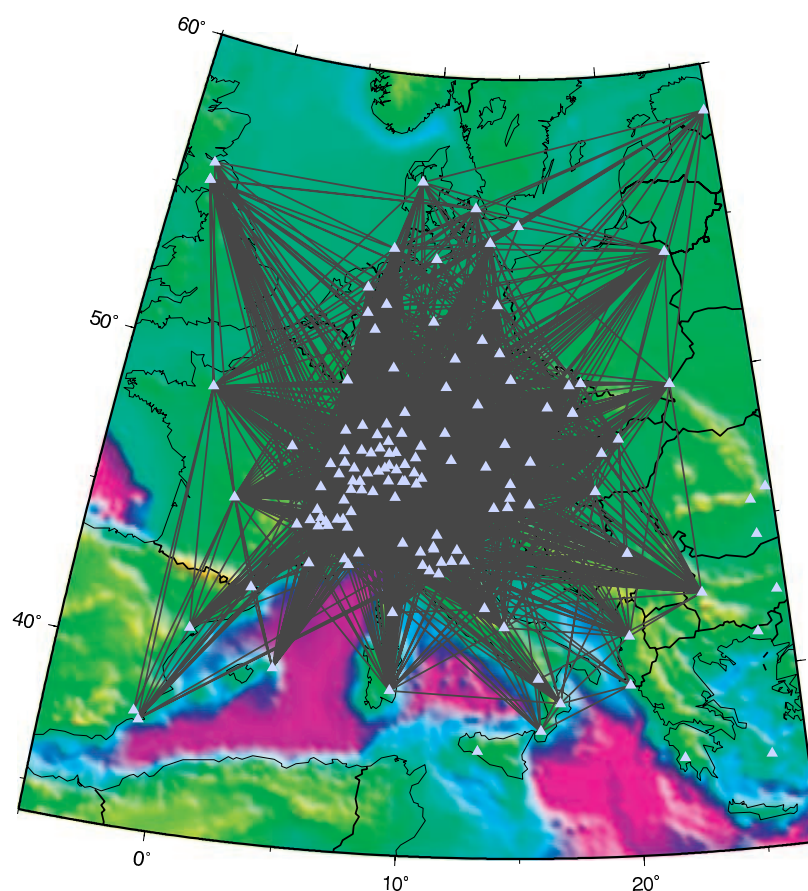


Figure 2 The 3000 selected paths where 16 s Rayleigh wave group velocity measurements were obtained from cross-correlations of ambient seismic noise. White triangles show stations used for this study.

We used one year of continuous records from October 2004 to October 2005 from 150 3-components broadband European stations. Our aim is to focus on the Alps, where we have a particularly high density of stations (Fig. 2). All the records are processed day per day. First the data are decimated to 1 Hz and corrected for the instrumental response. North and east horizontal components are rotated to obtain radial and transverse components with respect to the inter-station azimuth. The records are then band-pass filtered and their spectrum whitened between 5 and 150 s. We correlated signals recorded on the components that correspond to the non-zero terms of the theoretical elastic Green's tensor (ZZ, ZR, RZ, RR, and TT, due to symmetry considerations). Subsequently, correlations of one-day records are stacked. This is approximately equivalent to cross-correlating directly the whole year of records.

Rayleigh and Love waves dispersion curves are evaluated from the emerging Green's function using frequency-time anal-

ysis (Levshin *et al.* 1989; Ritzwoller and Levshin 1998) for the 17,000 inter-station paths. For each path we obtain eight evaluations of the Rayleigh-wave dispersion curves by considering four components of the correlation tensor (ZZ, RR, RZ and ZR) and both the positive and the negative part of the noise correlation function. Similarly, we obtain two estimates of the Love-wave dispersion curves from positive and negative parts of TT correlations.

We reject waveforms 1) with signal-to-noise ratio (ratio between Rayleigh wave's amplitude and noise variance after it) lower than seven; 2) with group velocities measured on the positive and negative correlation time differing by more than 5 percent; and 3) with paths shorter than two wavelengths at the selected period for the group velocity map. This results in about 3,500 paths over the initial 11,000 inter-stations paths at 16 s (Fig. 2). We then apply a tomographic inversion following Barmin *et al.* (2001) to this data set to obtain group velocity maps on $100 \times 100 = 10,000$ cells of 25×25 km

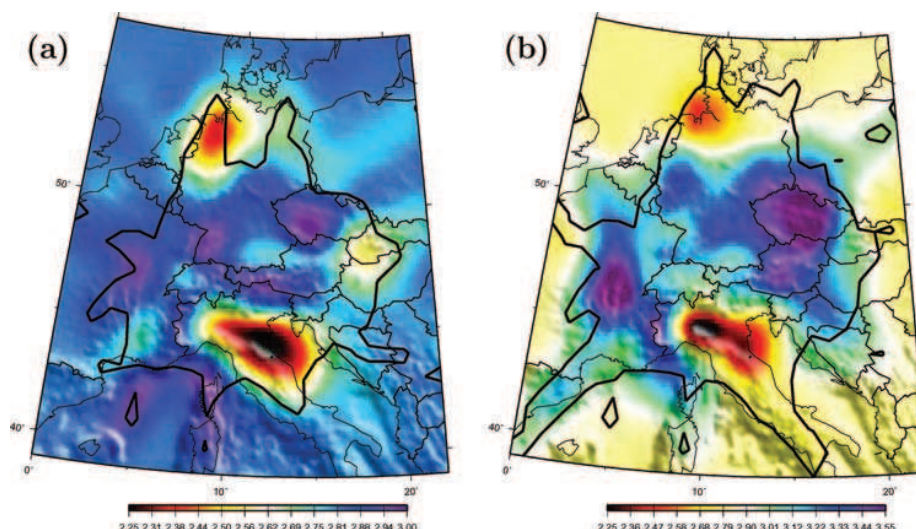


Figure 3 Rayleigh (a) and Love (b) wave group velocity maps at 16 s period constructed from 3,500 and 4,400 inter-station cross-correlations, respectively. Black thick line delimits the area where there are more than 10 paths per 25×25 km cell.

across Europe (Fig. 3). Several geological features can be seen on those maps. Low velocity anomalies are associated with sedimentary basins, such as the Po basin (Northern Italy), the North Sea basin and the Pannonian basin (Slovakia and Hungary). Both Rayleigh and Love waves exhibit smaller values below the molassic sediments (Southern Germany and Austria) than in the surrounding area. Close to the French-Italian border, one can notice a high-velocity anomaly corresponding to the Ivrea body, an intracrustal high-velocity and high-density zone within the Adriatic plate. The final resolution is good enough to see the contrast in Rayleigh wave velocity between the sedimentary (north-west) and the mountainous (south) part of Switzerland.

It is not possible to compare directly these group-velocity maps with maps obtained by active methods: practically, below 20 s of period, attenuation as well as scattering in the medium are too strong, preventing accurate measurement of surface-wave velocity from earthquakes or any active source. At these periods, the number of usable paths is thus too low to build any group-velocity map.

It is however possible to compare dispersion curves measured from passive and active methods in some other cases. Shapiro and Campillo (2004) measured dispersion curves using noise cross-correlation computed between pairs of stations separated by distances ranging from one hundreds to two thousand kilometres. At periods below 60 s, the resulting dispersion curves are in good agreement with those predicted by global group-velocity maps from Ritzwoller *et al.* (2002). In Southern California, Shapiro *et al.* (2005) compared records of

an earthquakes which occurred close to a station and recorded at two other stations, with noise cross-correlation signal computed on the same path. The measured arrival time was identical at the period ranging from 5 to 20 s. These results show the robustness of measurements performed using noise cross-correlations.

The Alpine region has intensively been studied using controlled source and earthquake tomography. These studies gave precise insight about the crustal and upper mantle structure (Marchant and Stampfli 1996; Waldhauser *et al.* 1998, 2002; Bleibinhaus and Gebrande 2005 and references therein). However, using seismic noise instead has several advantages. The final resolution depends mostly on the density of stations and is not limited by the available sources. This makes it possible to obtain high-resolution group velocity maps that cover large regions, whereas controlled sources can only be used for small areas. Surface wave tomography using earthquakes records only provides group velocity maps at periods above 20 s, since all the high-frequency information is lost due to attenuation in the medium.

3D S-WAVE TOMOGRAPHY OF THE PITON DE LA FOURNAISE VOLCANO

The same seismic noise cross-correlation technique can be applied to study more complex structures. In this section we present the 3D velocity model of a volcano obtained using only noise records.

Eighteen months (July 1999 to December 2000) of continuous seismic noise recorded at 21 vertical short period stations were collected by the Observatoire Volcanologique du Piton de la Fournaise (Fig. 4a). An example of a noise record at one of the stations (ANR) is shown in Fig. 4b. All noise records are first band passed between 1 to 5 s and their spectral amplitudes whitened in order to avoid strong dominant spectral peaks in the background noise.

For each available station pair, the one-bit noise correlation function is computed day per day. Some of the noise correlation functions are rejected upon a signal-to-noise ratio criterion (i.e. if the energy of the Rayleigh arrival is lower than 1.5 times the energy of the noise). The remaining traces are stacked over 18 months. For each path, group velocity dispersion curves are estimated using a frequency-time analysis (Levshin *et al.* 1989; Ritzwoller and Levshin 1998). We manually select dispersion curves according to group velocity limits and for station-to-station distances longer than one wavelength. We finally obtain 75 reliable dispersion curves from which group velocities are extracted for periods equal to 2, 2.5, 3, 3.5, 4 and 4.5 s.

The 2D Rayleigh wave group velocity maps are obtained from tomographic inversion of the arrival-time measurements at each period using the algorithm described by Barmin *et al.* (2001). Our 2D models involve $22 \times 28 = 616 \times 1$ km cells. Because of the sparse ray coverage and the low resolution of the data set, we choose to apply a strong smoothing to the tomographic inversion which results in a ~ 4 -km spatial resolution. The inversion results are thus robust and show a moderate variance reduction varying from 38 to 18% with increasing periods (from 2 to 4.5 s).

Dispersion curves are then extracted from the Rayleigh-wave group-velocity maps for every model cell. We fit these curves by polynomial functions in a least-squares sense and invert them using a Monte-Carlo algorithm (Shapiro *et al.* 1997), the synthetic dispersion curve being calculated using a method due to Herrmann (1987). We thus obtain a S-wave velocity-versus-depth profile for each cell. We present six horizontal slices as well as a 3D view of the 3D smoothed model in Figure 5. The results clearly show the presence of a high-velocity anomaly which moves westward with depth (+1.3 to -1.1 km above sea level). This structure is surrounded by a low-velocity ring interpreted as effusive products associated with the construction of the Piton de la Fournaise volcano on the flank of the older Piton des Neiges volcano. This high-velocity anomaly has also been detected by a previous earthquake and active P-wave tomography on the Piton de la Fournaise volcano (Lankar 1997). Recent works also imaged

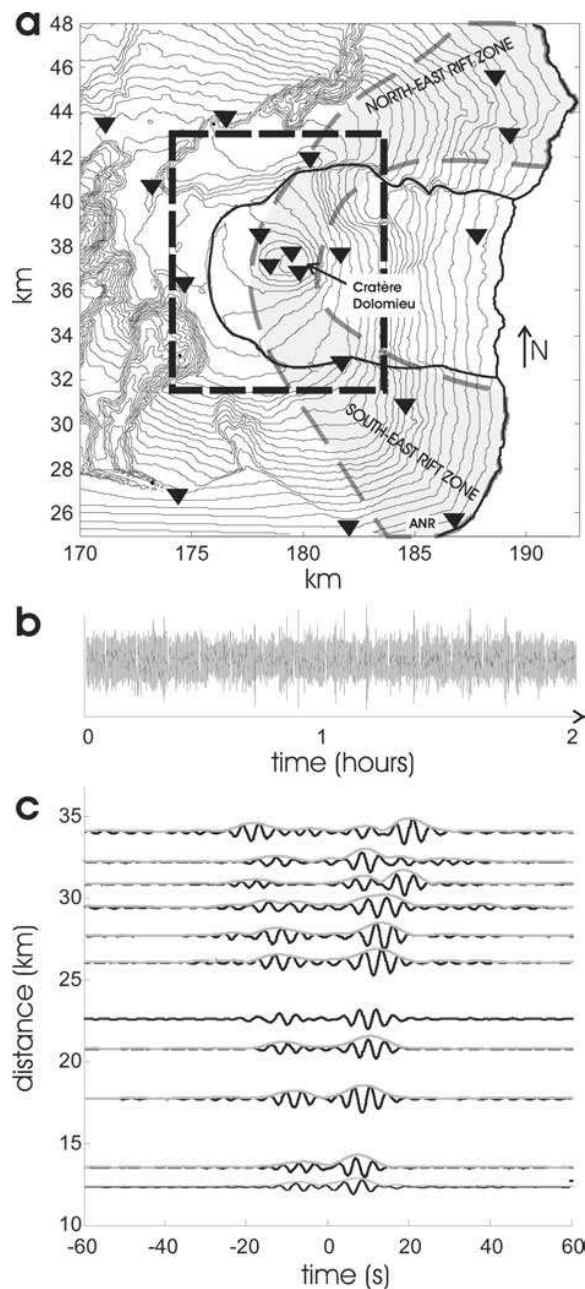


Figure 4 a. Map of the Piton de la Fournaise volcano. Seismic stations are represented as inverted triangles. The gray zone indicates the limits of the rift zone. The thin dashed rectangle corresponds to the limits of the presented tomographic images. Geographic coordinates are Gauss-Laborde kilometric coordinates (Transverse Mercator). Contour lines are spaced every 100 m. b. Two hours of ambient seismic noise at station ANR. c. Causal and acausal reconstructed Rayleigh waves (positive and negative times, dominant period 4 s) between station RMR (not shown on the map) and the rest of the network. The trace envelopes are represented as thin gray curves.

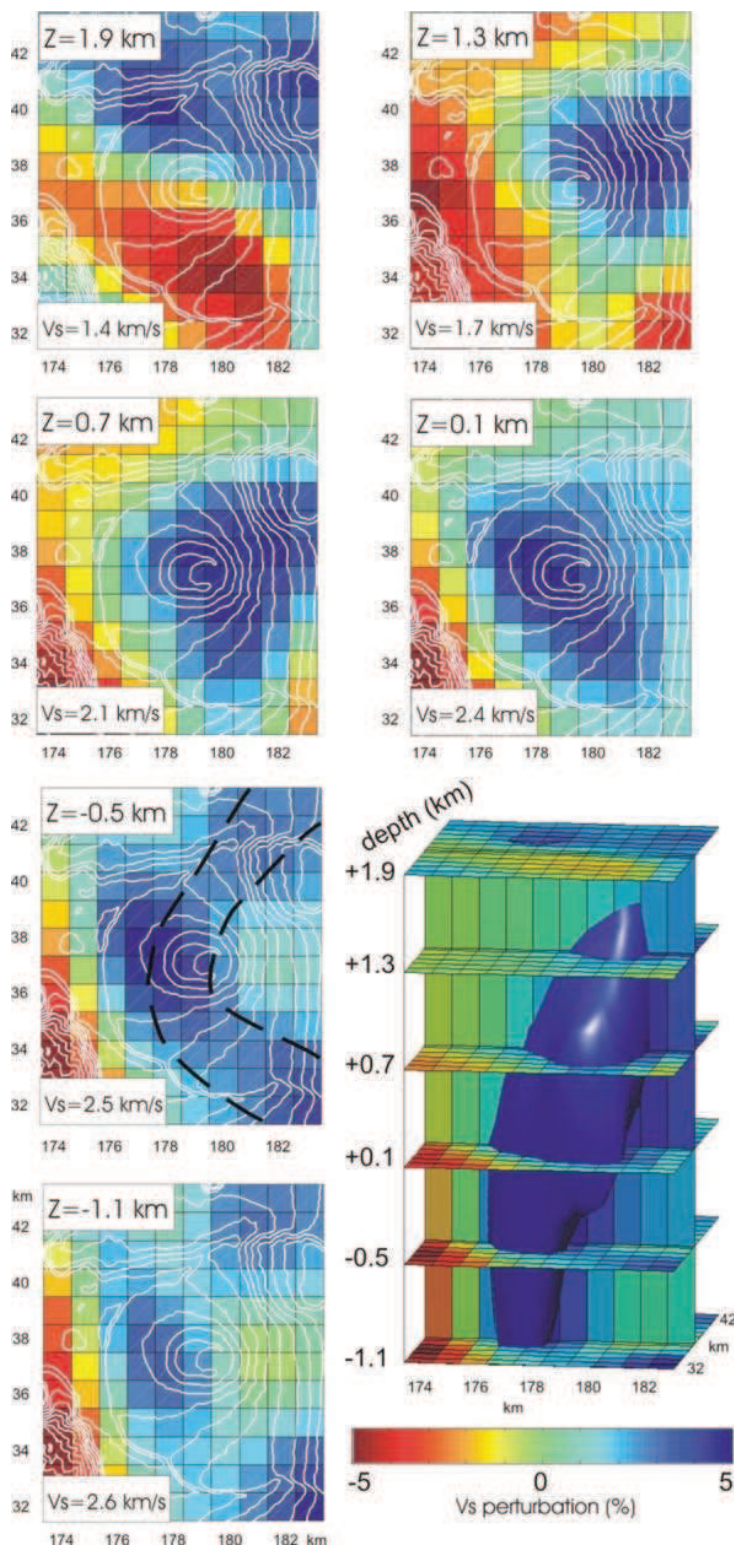


Figure 5 3D S-wave velocity model. We show 6 horizontal slices extracted from the 3D model at different depths. Average S-wave velocity is shown in white boxes on the bottom of corresponding slices. Black dashed line at depth -0.5 km shows the limits of the rift zone. We also plot a 3D view of the model. The 3D blue patch delimits the iso-velocity perturbation surface corresponding to a 2.5% velocity perturbation.

the presence of a high-velocity chimney on different volcanoes (Laigle *et al.* 2000; Tanaka *et al.* 2002; Zollo *et al.* 2002; Sherburn *et al.* 2006; Patan *et al.* 2006). We interpret this anomaly as solidified intrusive magma bodies. The high-velocity anomaly is also well correlated with the rift zone at sea level ($z = -0.5$ km). Imaging these intrusive bodies is of particular interest because the magma paths are usually believed to follow their geometry (Laigle *et al.* 2000; Battaglia *et al.* 2005). Furthermore, other studies showed that a few months of seismic noise data will yield similar three-dimensional results to that obtained from 18 months' data (Brengruier *et al.* 2008). We also achieved a preliminary study on the temporal variations of the reconstructed Green's functions showing that we could detect relative velocity variations of less than 0.1% with a temporal resolution of one day.

P-WAVES EXTRACTION FROM SEISMIC NOISE CROSS-CORRELATION

A main issue in the convergence of the correlation process to the transfer function is the influence of variations in the temporal and spatial distribution of the noise sources. From the temporal point of view, the noise spectrum defines the frequency bandwidth over which the impulse response can be retrieved. When receivers are widely separated, the coherent propagating noise must have sufficient amplitude to be recorded on both receivers despite geometrical spreading and attenuation. This explains why the slowly-attenuated Rayleigh waves have dominated the impulse response obtained so far from correlations of seismic noise.

In this section, we present results from the correlation function of seismic noise recordings among pairs of stations in the dense Parkfield network, California. When performed on many station pairs at short ranges, the noise correlation function demonstrates the presence of both a P-wave and a Rayleigh wave in the noise correlation function.

We processed data recorded on the dense temporary seismic network installed in the Parkfield area between July 2001 and October 2002. One month of seismic noise recordings were cross-correlated between each pair of 30 broadband 3-component seismic stations located in an 11-km square (Fig. 6, Thurber *et al.* 2004). We used only the vertical component. This network has extensively been used to monitor and image the San Andreas Fault using both man-made explosions and earthquakes as part of the San Andreas Fault Observatory at Depth (SAFOD) project. Inversion results have confirmed the spatial heterogeneity of P-wave velocity across the Fault

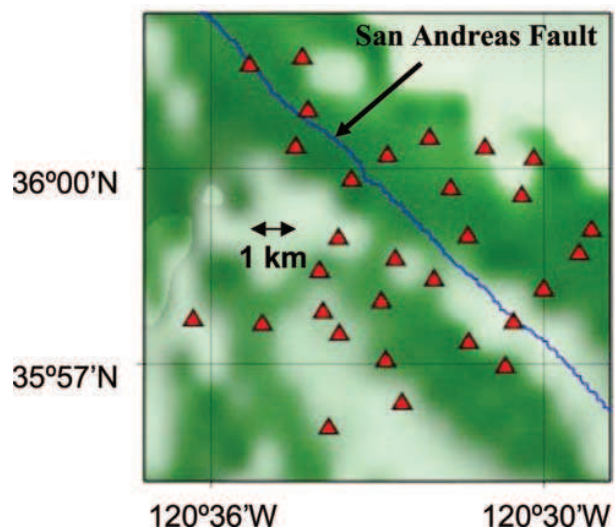


Figure 6 Topographic map of the Parkfield area (an 11-km large square) showing stations (triangles) and SAF (blue).

up to 6 km in depth (Ben-Zion and Malin 1991; Catchings *et al.* 2002; Hole *et al.* 2006).

At first, frequency-incoherent beamforming is performed using the $N = 30$ stations of the network to determine the average velocity c and direction θ of the seismic noise (Fig. 7). Beamforming is performed in two frequency bands of interest [0.2–0.5 Hz] and [0.6–1.3 Hz], on data segments of one day of seismic noise, which were recorded in February 2002, as:

$$B(\theta, c) = \frac{1}{\delta\omega} \int_{\omega_c - \delta\omega/2}^{\omega_c + \delta\omega/2} \sum_{i=1}^N \hat{S}_i(\omega) \exp\left(i \frac{\omega}{c} (x_i \sin \theta + y_i \cos \theta)\right) d\omega$$

where ω_c is the central noise frequency and $\delta\omega$ the frequency bandwidth, $\hat{S}_i(\omega)$ is the complex Fourier component at frequency ω of the noise signal $S_i(t)$ recorded on the i th seismic station ($i \in [1, N]$), and (x_i, y_i) are the longitude/latitude coordinates of station i . Working with a dense seismic network having a small coverage area allows determination of an average apparent incoming velocity for this zone using plane wave beamforming. In both of the frequency bands, the noise field clearly originates from the Pacific Ocean with an incident direction $\theta_0 = 55^\circ$ on the Parkfield network. On the other hand, the beamformer in the [0.2–0.5 Hz] band exhibits an apparent velocity of 2.8 km/s compatible with a Rayleigh wave (Fig. 7a), while the beamformer in the [0.6–1.3 Hz] band shows an apparent velocity of 5 km/s (Fig. 7b).

Since Fig. 7 reveals a strong directivity in the seismic noise, only station pairs aligned with the noise main direction θ_0 are

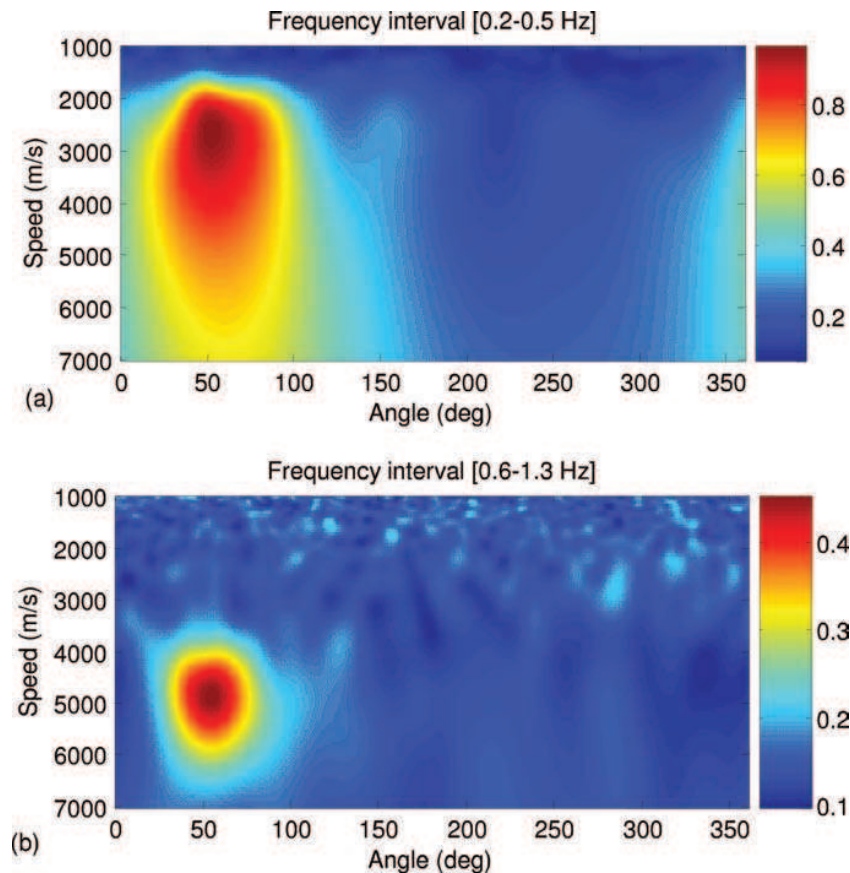


Figure 7 Angular-speed distribution of pre-processed incoming noise on the Parkfield network averaged over one month. Plane wave beam-forming is summed incoherently over 100 frequencies from (a) 0.2 to 0.5 Hz and (b) 0.6 to 1.3 Hz. The x-axis corresponds to noise directivity θ , north is 0° .

chosen to compute the point-to-point noise correlation function. Restricting the cross-correlations to these pairs ensures that 1) the travel time of the main peak of the noise correlation function is not biased and corresponds to the actual travel time of the Green's function between the stations (Snieder 2004; Roux *et al.* 2005b) and 2) the signal-to-noise ratio of the noise correlation function is maximized. Practically speaking, 145 station pairs are selected in the Parkfield network whose angles θ_{ij} are included in a directivity angle $\delta\theta = |\theta_{ij} - \theta_0| \leq \sqrt{\frac{c}{R_{ij} \omega_c}}$ dependent on the distance R_{ij} between stations and the frequencies characteristics of the seismic noise field (Roux and Kuperman 2004).

The noise correlation function is computed for each selected station pair as in equation 3 of Section 2, and averaged over 30 days to further increase the signal-to-noise ratio. Figure 8a is a display of the noise correlation functions obtained for the 145 selected pairs sorted by ascending offset R_{ij} . A prop-

agating wavefront clearly appears at high frequency (Fig. 8b) which corresponds to a ~ 5 km/s velocity wave. A polarization study between the Z-R and Z-Z components of the correlation tensor confirmed the P-wave nature of this wavefront (Roux *et al.* 2005a).

Going back to Fig. 7(b), we note that the apparent velocity of the P-wave corresponds to the P-velocity at the turning point. Recent inversions of the P-wave velocity profile reveal a strong velocity gradient at the surface, the 5 km/s speed being reached at no more than 1.5 km in depth on the west side of the San Andreas fault. This confirms that the noise sources that excite P-waves are local and cannot be confused with deep incident waves that would hit the seismic array with a much higher apparent velocity. One hypothesis is that P-waves are locally generated by conversion of incident Rayleigh waves coming from the Pacific by local heterogeneities at the Earth's surface or subsurface.

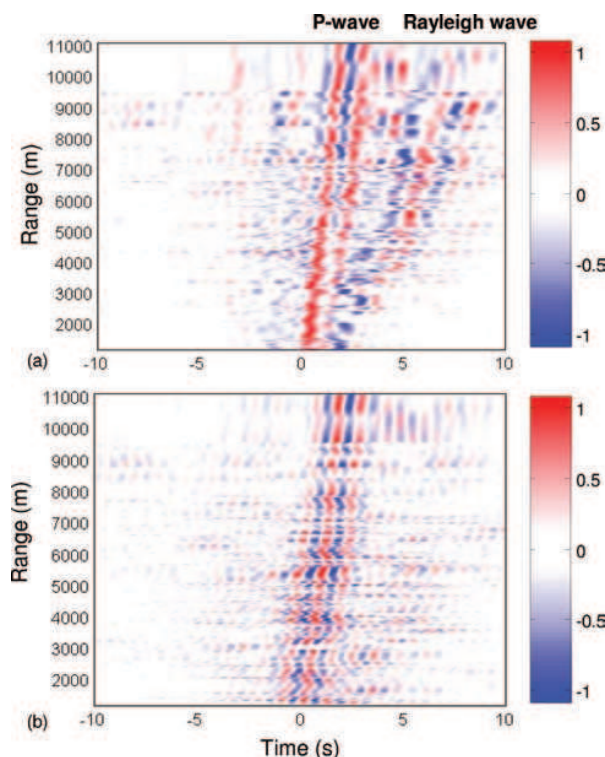


Figure 8 Range-time representation of the Z-Z component of the noise correlation tensor averaged over one month in two frequency bands (a) [0.1–1.3 Hz], and (b) [0.6–1.3 Hz]. Each plot has been normalized by its own maximum.

SMALL SCALE GEOPHYSICS USING SURFACE WAVES EXTRACTED FROM NOISE CROSS-CORRELATION

In this section, we achieve an experimental demonstration of the correlation process of controlled noise sources at the metre scale using a linear array of accelerometers. It is known that the cross-correlation of seismic noise between two receivers converges towards the Green's function when noise is equidistributed in azimuth. The same result can be obtained with directional noise if noise sources are located in the end-fire lobes centered along the array line direction (Roux and Kuperman 2004). When noise is not isotropic and noise sources are not adequately located with respect to the receivers array, 'controlled' noise sources can be used to produce appropriate wavefield satisfying the end-fire lobe criterion. This is particularly useful at small scales and high frequencies where local sources strongly contribute to the noise wavefield.

A 14-metre long line of 8 evenly-spaced vertical accelerometers has been used to record human steps. We walked in the alignment of the accelerometers line, 5 times one minutes

on each side, from 0 to about 30 metres away from the accelerometer array. The experimental setup was designed to be versatile: 1) the array configuration may include up to 16 one-component seismic stations; 2) these seismic sensors could be accelerometers or geophones depending on the expected frequency bandwidth; 3) the array length is adjustable to the surface wave wavelength.

The main advantage of this system is to be easy and fast to setup. Our ambition was to achieve a complete deployment, acquisition and processing in approximately 30 minutes. The system design makes it very convenient for local and near surface measurements.

The energy spectrum of the recorded steps spreads up to 150 Hz. Given the frequency response of the accelerometers and the spatial extension of the array, a frequency interval ranging from 10 to 100 Hz was selected for the analysis. Since the frequency spectrum of the steps is not flat in this frequency interval, and as correlating is mathematically equivalent to a spectrum product, only the most energizing frequencies will emerge in the correlation signal. To enlarge the effective frequency bandwidth, the spectrum of the records is equalized in the selected frequency interval [10–100 Hz] before the correlation process.

To check the robustness of the correlation process, five one-minute long records were separately correlated for each accelerometer pair. The five time-domain correlations superimpose in phase, leading to the conclusion that correlation is robust and does not depend on the way we walked. As these correlation signals superimpose, they are stacked to increase the signal-to-noise ratio. The advantage of the correlation process is then to perform an ensemble average over the 'controlled' noise sources without the need of synchronization. The superposition of the correlations of one minute long signals is thus just a verification of the repeatability of the steps. Stacking five correlations of one minute long records is equivalent to correlating directly a 5-minute long signal.

To obtain a seismic section from the correlation process, signals are cross-correlated by the accelerometers located at the extremity of the line array. Taking one or the other of the accelerometers as the reference signal does not modify the seismic section. This shows that seismic propagation from left to right is identical to propagation from right to left on the 14-m long seismic array. The medium can then be assumed 1D in the frequency bandwidth of the recordings. The 1D argument can be pushed even further. Two receiver pairs separated by the same range are stacked since propagation does not depend on the pair location but only on the offset between receivers.

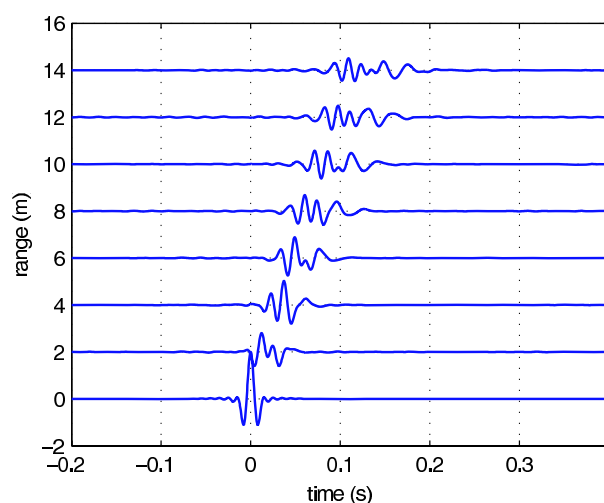


Figure 9 Seismic section obtained from the correlation process after all averaging operations. The signal-to-noise ratio is above 30 dB for each trace. Both phase dispersion and attenuation are observed. The seismic propagation reveals the presence of two surface waves with group velocity of about 90 and 120 m/s.

Figure 9 shows the seismic section obtained after the complete spatial and temporal stacking. After a 10-minutes total recording, signal-to-noise ratio is above 30 dB. Both phase dispersion and attenuation during propagation is retrieved. The final seismic section clearly reveals the presence of two surface waves, with mean group velocities of about 90 and 120 m/s. Those low group velocities are good indications of two Rayleigh modes.

We wish to insist on the fact that this section was obtained from 10 minutes of unsynchronized human steps only, which makes it almost a passive method. To obtain the same result with usual active seismic techniques, much more time would have been needed to synchronize numerous sledgehammer blows. The 'passive' method presented here is thus 1) easy to implement, as there is a large flexibility in the array configuration, 2) fast, as it takes only about 30 minutes to complete the array deployment and the recording, and 3) simple, as there is no synchronization task and processing is performed in real time.

A frequency-wavenumber (F-K) transform was applied to the seismic section in Fig. 9 to identify each of the surface waves (Fig. 10). As the accelerometers are evenly spaced by a distance $d = 2$ m, the largest wavenumber before aliasing is $2\pi/d = \pi$. In Fig. 10, this value of k corresponds to the green line. Higher k 's are wrapped, and appear as low wavenumbers. In this simple case, the wavenumber spectrum can be extended by unwrapping the k axis.

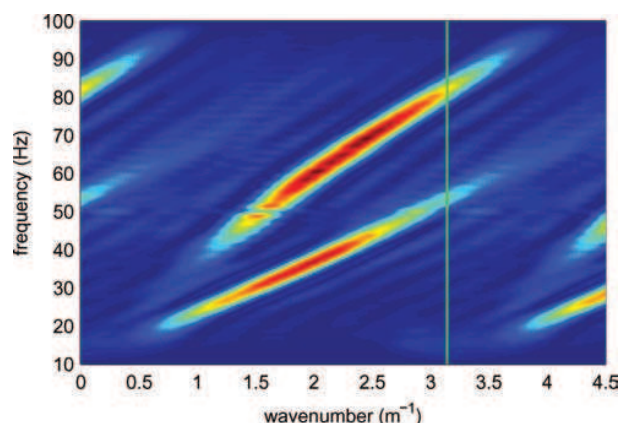


Figure 10 Frequency wavenumber (F-K) transform of the seismic section obtained in Fig. 9. The largest measurable wavenumber according to the Shannon criterion is $2\pi/d$ with $d = 2$ m (green line). The aliasing in the F-K diagram is resolved by unwrapping the k axis. The shape of the two modes on the F-K diagram reveals dispersive modes.

From the F-K diagram in Fig. 10, modes are separated and their phase velocity dispersion curves are extracted. Those surface-wave dispersion curves are the starting point for a surface wave inversion to retrieve the local velocity versus depth profile of the medium.

PASSIVE CORRELATION IMAGING OF A BURIED SCATTERER

Up till now, most of geophysical applications of passive imaging with ambient-noise cross-correlation have been used to reconstruct direct arrivals of Rayleigh or P-waves. Reconstructing other features of the Green's function, like the reflections following direct waves, is harder: the reflections are weaker and the propagation is fully 3D. Nevertheless, passively imaging a scatterer would form a major application to prospecting and certainly deserves attention. In order to test the feasibility of passively imaging a buried scatterer, we set a controlled ultrasonic experiment in the laboratory. We believe the principles presented here also apply to ambient seismic noise.

To mimic micro-seismic vibrations, we use a highly reverberant body excited by a series of sources (see Fig. 11). A 12 mm diameter cylindrical hole was drilled through an aluminium block of dimensions 125 mm \times 125 mm \times 90 mm. The hole is 25 mm beneath the surface. To excite elastic waves, we employ a laser mounted on a step motor. For a complete description of the experimental set-up, please refer to Larose *et al.* (2006b). The laser emits mainly shear waves (Mason and Thurston 1988) (see directivity in Fig. 11). The resulting

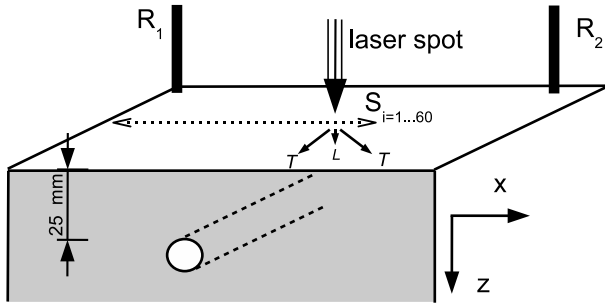


Figure 11 Set-up of the ultrasonic experiment. The source scans the surface with 1 mm steps along a line of 60 mm. The directivity (Mason and Thurston 1988) of one laser shot is shown for shear (transverse T) waves, and weaker compressional (longitudinal L) waves.

wave field is recorded by two pin transducers located at \vec{r}_1 and \vec{r}_2 . The laser triggers the waveform acquisition. For each position of the source \vec{r}_s^i and receiver k , the record is noted $S_{ik}(t) = G(t, \vec{r}_k, \vec{r}_s^i) \otimes R_k(t)$ where G is the elastic Green's function, \otimes is convolution and $R_k(t)$ is the transfer function of the receiver k . Each record is filtered in the [0.05–0.9 MHz] frequency band, where the absorption time of the block is about 30 ms. Diffuse field decay is therefore slow enough to permit record lengths greater than 100 ms, which represents thousands of reverberations within the cavity. After each acquisition, the laser is moved to another position. 1 mm steps are used to mimic a linear array of 60 points. By reciprocity, the sources and receivers can be interchanged. Our experimental set-up is therefore analogous to a conventional seismic experiment where a linear array of 60 geophones would sense the seismic diffuse wavefield generated by two distant sources. The Green's function between any couple of points (\vec{r}_s^i, \vec{r}_s^j) of the array is recovered by processing the following time-correlation:

$$\begin{aligned} C_{ij}^k(\tau) &= \int S_{ik}(t) S_{jk}(t + \tau) dt \\ &= G(t, \vec{r}_k, \vec{r}_s^i) \times G(t, \vec{r}_k, \vec{r}_s^j) \otimes R_k(t) \otimes R_k(-t) \end{aligned}$$

To remove the receiver functions R_k , we deconvolve the cross-correlations by the averaged auto-correlations $\langle C_{ii}^k(\tau) \rangle_i \approx R_k(t) \otimes R_k(-t)$. This procedure has the additional virtue of removing contaminations $G(t, \vec{r}_k, \vec{r}_k)$ or *ghosts*, which Derode *et al.* (2003b), Weaver and Lobkis (2006) related to the environment of the receivers R . Then these correlations are averaged over the available distant sources $\vec{r}_{k=1,2}$ to obtain $C_{ij}(\tau)$. As noted by several authors this correlation is essentially the Green's function $G(t, \vec{r}_s^i, \vec{r}_s^j)$ and therefore should contain the deterministic signature of the isolated scatterer.

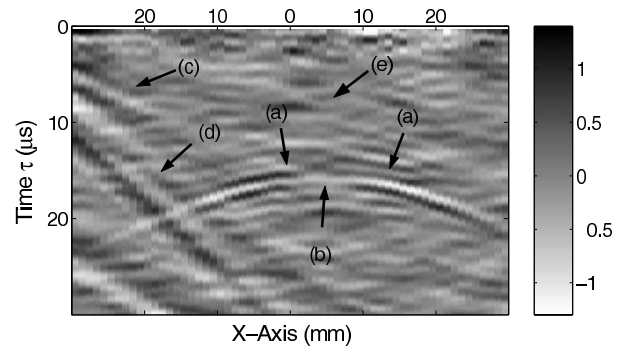


Figure 12 Time-distance wavefield (linear scale, arbitrary unit). Each autocorrelation $C_{ii}(\tau)$ is plotted for different position i along the X-axis, and time τ . Position 0 marks the center of the array; (a-e) are different reflections (see the text).

In Fig. 12 we display the time-distance wavefield obtained for all the available autocorrelations $C_{ii}(\tau)$. This autocorrelation is the field sensed in i if i were both source and receiver. Position 0 marks the centre of the array. The hyperbolic feature is the signature of the buried scatterer: the wave labeled (a) is the wave reflected by the top of the cylindrical hole placed at $z = 25$ mm and $x = 4$ mm. The arrival times along the array correspond to a shear (transverse) wave ($v_T = 3.1$ mm/ μ s). (c) is a compressional-to-Rayleigh reflected by the lateral edge of the cavity, and (d) is a Rayleigh-to-Rayleigh reflected by the same edge. The shear wave directivity of the laser generation is clearly visible in the null at (b). Because longitudinal wave generation is much weaker than that of shear waves (Mason and Thurston 1988), its reflection (e) is hardly visible.

A noteworthy point is that the passive reconstruction of any $G(t, \vec{r}_s^i, \vec{r}_s^j)$ remains imperfect. The averaging used to construct the correlation is finite, leaving visible fluctuations in Fig. 12. To improve the quality of this image, one could increase the record length, or employ additional receivers \vec{r}_k (Weaver and Lobkis 2005b). Alternatively, we could perform beamforming in order to take advantage of all the $C_{i \neq j}$ cross-correlations. Beamforming is a standard procedure to obtain medical or seismic (migrated) image. The new point is that here the impulse responses $C_{i \neq j}$ are obtained *passively*.

The image we now process is a 2D image of the reflectivity of the medium. The first step is to apply beamforming to the forward propagation to focus the wave on any point (x, y) in the medium. This is achieved by summing the time-delayed impulse responses $C_{ij}(\tau)$. The same beamforming technique is then applied to the wave back-propagation (from the focal

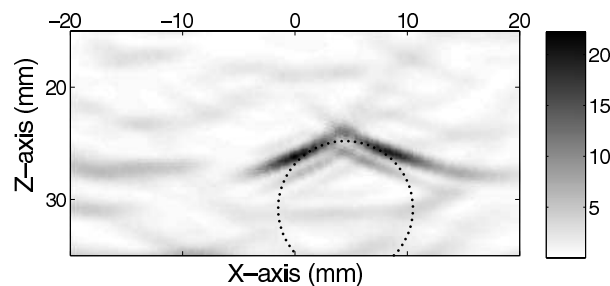


Figure 13 Reflectivity (linear scale, arbitrary unit) of the aluminium block as probed by bulk shear waves. The array of 60 laser sources is at $z = 0$. Black indicates a high reflectivity. The top of the hole is clearly visible. The actual position of the cylinder is displayed in the dotted line.

point to the receivers), the reflectivity ϕ of the medium is then:

$$\phi(x, y) = \left[\sum_{ij} C_{ij}(\tau_i + \tau_j) \right]^2$$

where $\tau_i = \frac{1}{v_T} \sqrt{(x - x_i)^2 + z^2}$ and v_T is the shear wave velocity. The 2D reflectivity map of the medium is displayed in Fig. 13. The top of the reflector is clearly visible. Because of the finite size of the linear array, the sides and bottom of the cylindrical hole cannot be imaged. The shear wave directivity of the laser spot induces a preferential reflection for oblique incidences; additionally, the images in Figs 12 and 13 show a null at apex. Speckle fluctuations are noticeable making these figures a little more noisy than the ones obtained with Rayleigh waves. This is expected since the field at the free surface is dominated by Rayleigh waves.

To conclude this part, we have shown here the feasibility of imaging small details of the medium (like a buried isolated scatterer) by means of the passive time-correlation technique. By reciprocity this experimental set-up is analogous to an array of seismic geophones sensing the diffuse wavefield originating from distant sources. We therefore believe this technique could be transposed and applied to geophysical prospecting, as well as to medical imaging. The use of fully developed diffuse field in a closed cavity is not a rigorous requirement for this imaging technique. It could in principle be replaced by any other diffuse field, like diffuse waves in an open medium, or ambient noise.

CONCLUSION

In this paper, we theoretically reviewed how and under which assumptions cross-correlation of noise recorded at two sensors yields the Green's function between them. This property is based on equipartition of the wavefield that can be provided

either by an appropriate sources distribution or by wave scattering in the medium. Any diffuse field, like diffuse waves in an open medium, or ambient noise may be used to reconstruct the Green's function between two points.

We experimentally showed the feasibility of passive imaging using noise cross-correlation. This technique bypasses the usual shortcomings encountered in active imaging, especially concerning the requirements about sources (strength, location, occurrence, etc). Application to ambient seismic noise is particularly promising for improving images of the Earth as the number of usable ray paths for tomography is directly linked to the number of recording stations. In more complex structures like volcanoes, this technique was validated as the S-wave velocity model obtained presents the same anomalies as in active measurements. Nevertheless, in a context where standard active methods are hampered by irregular sources distribution, the possibility of using noise records is particularly interesting. These results demonstrate the possibility to achieve surface-wave tomography from noise cross-correlation. Results from the Parkfield area show that P-waves are present in the correlation on small-scale seismic networks, and could be used for body wave tomography.

At smaller scales, cross-correlation imaging techniques brings a new way to achieve seismic experiments that is faster and easier to implement than usual active methods. The recorded wavefield can be produced by active sources adequately located or using scattering to produce a diffuse wavefield. We showed that late arrivals, like reflections produced by buried objects, can be retrieved. This passive imaging of scatterers would be a major application of noise cross-correlation in geophysical prospecting. This technique is also promising for geophysical surveys, as seismic noise is a reproducible, stationary in time, natural source, that could also be used to give new insights into 4D seismic exploration.

REFERENCES

- Aki K. 1957. Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bulletin of the Earthquake Research Institute* 35, 415–456.
- Barmin M.P., Ritzwoller M.H. and Levshin A.L. 2001. A fast and reliable method for surface wave tomography. *Pure and Applied Geophysics* 158, 1351–1375.
- Battaglia J., Ferrazzini V., Staudacher T., Aki K. and Cheminée J.-L. 2005. Pre-eruptive migration of earthquakes at the Piton de la Fournaise volcano (Réunion Island). *Geophysical Journal International* 161, 449–458.
- Ben-Zion Y. and Malin P. 1991. San Andreas fault zone head waves near Parkfield, California. *Science* 251, 1592–1594.

- Bleibinhaus F. and Gebrande H. 2005. Crustal structure of the Eastern Alps along the TRANSALP profile from wide-angle seismic tomography. *Tectonophysics* **414**, 51–69.
- Brenguier F., Shapiro N.M., Campillo M., Nercissian A. and Ferrazzini V. 2007. 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlation. *Geophysical Research Letters* **34**, L02305.
- Brenguier F., Shapiro N.M., Campillo M., Ferrazzini Z., Duputel Z., Coutant O. and Nercissian A. 2008. Toward forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, in press.
- Campillo M. 2006. Phase and correlation in ‘random’ seismic fields and the reconstruction of the Green function. *Pure and Applied Geophysics* **163**, 475–502.
- Campillo M. and Paul A. 2003. Long-range correlations in the diffuse seismic coda. *Science* **299**, 547–549.
- Catchings R.D., Rymer M.J., Goldman M.R., Hole J.A., Huggins R. and Lippus C. 2002. High-resolution seismic velocities and shallow structure of the San Andreas fault zone at Middle Mountain, Parkfield, California. *Bulletin of the Seismological Society of America* **92**, 2493–2503.
- Clairbout J.F. 1968. Synthesis of a layered medium from its acoustics transmission response. *Geophysics* **33**, 264–269.
- Colin de Verdière Y. 2006a. Mathematical models for passive imaging I: general background. URL <http://fr.arxiv.org/abs/math-ph/0610043/>.
- Colin de Verdière Y. 2006b. Mathematical models for passive imaging II: effective hamiltonians associated to surface waves. URL <http://fr.arxiv.org/abs/math-ph/0610044/>.
- Derode A., Larose E., Campillo M. and Fink M. 2003a. How to estimate the Green’s function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Applied Physics Letters* **83**, 15, 3054–3056.
- Derode A., Larose E., Tanter M., de Rosny J., Tourin A., Campillo M. and Fink M. 2003b. Recovering the Green’s function from field-field correlations in an open scattering medium (L). *The Journal of the Acoustical Society of America* **113**, 2973–2976.
- Duvall T.L., Jefferies S.M., Harvey J.W. and Pomerantz M.A. 1993. Time distance helioseismology. *Nature* **362**, 430–432.
- Gilles P.M., Duvall T.L., Scherrer P.H. and Bogart R.S. 1997. A sub-surface flow of material from the Sun equator’s to its poles. *Nature* **390**, 52–54.
- Godin O.A. 2007. Emergence of the acoustic Green’s function from thermal noise. *The Journal of the Acoustical Society of America* **121**, EL96–EL102.
- Gouédard P., Roux P. and Campillo M. 30 August - 01 September 2006. Small scale geophysics inversion using surface waves extracted from noise cross-correlation. In *3rd International Symposium on the Effects of Surface Geology on Seismic Motion*.
- Herrmann R.B. 1987. Computer programs in seismology. URL <http://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html>.
- Hole J., Ryberg T., Fuis G., Bleibinhaus F. and Sharma A. 2006. Structure of the San Andreas fault zone at SAFOD from a seismic refraction survey. *Geophysical Research Letters* **33**, L07312.
- Kang T.-S. and Shin J.S. 2006. Surface-wave tomography from ambient seismic noise of accelerometer networks in southern Korea. *Geophysical Research Letters* **33**, L17303.
- Laigle M., Hirn A., Sapin M., Lépine J., Diaz J., Gallart J. and Nicolich R. 2000. Mount Etna dense array local earthquake P and S tomography and implications for volcanic plumbing. *Journal of Geophysical Research* **105**, 21.633–21.646.
- Lankar V. 1997. *Approches par tomographie sismique du Piton de la Fournaise, La Réunion*. Thèse de doctorat, Institut de Physique du Globe de Paris, France.
- Larose E., Khan A., Nakamura Y. and Campillo M. 2005. Lunar sub-surface investigated from correlation of seismic noise. *Geophysical Research Letters* **32**, L16201.
- Larose E., Margerin L., Derode A., Tiggelen B.V., Campillo M., Shapiro N.M., Paul A., Stehly L. and Tanter M. 2006a. Correlation of random wavefields: an interdisciplinary review. *Geophysics* **71**, 4, SI11–SI21.
- Larose E., Derode A., Campillo M. and Fink M. 2004. Imaging from one-bit correlation of wide-band diffuse wavefield. *Journal of Applied Physics* **95**, 8393–8399.
- Larose E., Lobkis O.I. and Weaver R.L. 2006b. Passive correlation imaging of a buried scatterer. *The Journal of the Acoustical Society of America* **119**, 3549–3552.
- Levshin A., Yanocskaya T.B., Lander A.V., Bukchin B.G., Barmin M.P., Ratnikova L.I. and Its E.N. 1989. *Seismic surface waves in a laterally inhomogeneous Earth*. Kluwer Academic Publishers.
- Lin F., Ritzwoller M.H., Townend J., Bannister S. and Savage M. 2007. Ambient noise Rayleigh wave tomography of New Zealand. *Geophysical Journal International* **170**, 649–666.
- Lobkis O.I. and Weaver R.L. 2001. On the emergence of the Green’s function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America* **110**, 3011–3017.
- Marchant R. and Stampfli G. 1996. Subduction of continental crust in the Western Alps. *Tectonophysics* **269**, 217–235.
- Mason W.P. and Thurston R.N. 1988. *Physical Acoustics*, vol. **18**. Academic Press.
- Patanè D., Barberi G., Cocina O., De Gori P. and Chiarabba C. 2006. Time-resolved seismic tomography detects magma intrusions at Mount Etna. *Science* **313**, 821–823.
- Ramm A.G. 1986. *Scattering by obstacles*. Kluwer Academic Publisher.
- Reed M. and Simon B. 1978. *Methods of modern mathematical physics I, II, III, IV*. Academic Press.
- Ritzwoller M. and Levshin A.L. 1998. Eurasian surface wave tomography: group velocities. *Journal of Geophysical Research* **103**, 4839–4878.
- Ritzwoller M.H., Shapiro N.M., Barmin M.P. and Levshin A.L. 2002. Global surface wave diffraction tomography. *Journal of Geophysical Research* **107**, B12, 2335.
- Roux P. and Kuperman W.A. 2004. Extracting coherent wavefronts from acoustic ambient noise in the ocean. *The Journal of the Acoustical Society of America* **116**, 4, 1995–2003.
- Roux P., Sabra K.G., Gerstoft P. and Kuperman W.A. 2005a. P-waves from cross-correlation of seismic noise. *Geophysical Research Letters* **32**, L19303.
- Roux P., Sabra K.G., Kuperman W.A. and Roux A. 2005b. Ambient noise cross correlation in free space: theoretical approach. *The Journal of the Acoustical Society of America* **117**, 1, 79–84.

- Sabra K.G., Gerstoft P., Roux P. and Kuperman W.A. 2005a. Surface wave tomography from microseisms in southern California. *Geophysical Research Letters* **32**, L14311.
- Sabra K.G., Roux P. and Kuperman W.A. 2005b. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide. *The Journal of the Acoustical Society of America* **117**, 1, 164–174.
- Sabra K.G., Roux P. and Kuperman W.A. 2005c. Emergence rate of the time-domain Green's function from the ambient noise cross-correlation. *The Journal of the Acoustical Society of America* **118**, 6, 3524–3531.
- Sánchez-Sesma F.J. and Campillo M. 2006. Retrieval of the Green function from cross correlation: the canonical elastic problem. *Bulletin of the Seismological Society of America* **96**, 1182–1191.
- Sánchez-Sesma F.J., Pérez-Ruiz J., Campillo M., and Luzón F. 2006. Elastodynamic 2-D Green function retrieval from cross-correlation: canonical inclusion problem. *Geophysical Research Letters* **33**, L13305.
- Sánchez-Sesma F.J., Pérez-Ruiz J., Luzón F., Campillo M. and Rodríguez-Castellano A. 2007. Diffuse fields in dynamic elasticity. *Wave Motion*, in press.
- Shapiro N.M. and Campillo M. 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters* **31**, L07614.
- Shapiro N.M., Campillo M., Paul A., Singh S.K., Jongmans D. and Sánchez-Sesma F.J. 1997. Surface-wave propagation across the Mexican Volcanic Belt and the origin of the long-period seismic-wave amplification in the Valley of Mexico. *Geophysical Journal International* **128**, 151–166.
- Shapiro N.M., Campillo M., Stehly L. and Ritzwoller M.H. march 2005. High-resolution surface wave tomography from ambient seismic noise. *Science* **307**, 1615–1618.
- Sherburn S., White R.S. and Chadwick M. 2006. Three-dimensional tomographic imaging of the Taranaki volcanoes, New Zealand. *Geophysical Journal International* **166**, 957–969.
- Snieder R. 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase. *Physical Review E* **69**, 4 Pt 2, 046610.
- Stehly L., Campillo M. and Shapiro N. 2006. A study of the seismic noise from its long range correlation properties. *Journal of Geophysical Research* **111**, B10306.
- Stehly L., Campillo M. and Shapiro N.M. 2007. Travel time measurements from noise correlation: stability and detection of instrumental time-shifts. *Geophysical Journal International* **171**, 223–230.
- Tanaka S., Hamaguchi H., Nishimura T., Yamawaki T., Ueki S., Nakamishi H., Tsutsui T., Miyamachi H., Matsuwo N., Oikawa J., Ohminato T., Miyaoka K., Onizawa S., Mori T. and Aizawa 2002. Three-dimensional P-wave velocity structure of Iwate volcano, Japan from active seismic survey. *Geophysical Research Letters* **29**, 59–62.
- Thurber C., Roecker S., Zhang H., Baher S., Ellsworth W. and Tanimoto T. 2004. Fine-scale structure of the San Andreas Fault Zone and location of the SAFOD target earthquakes (2004). *Geophysical Research Letters* **31**, L12S02.
- van Tiggelen B.A. 2003. Green function retrieval and time reversal in a disordered world. *Physical Review Letter* **91**, 24, 243904.
- Waldhauser F., Kissling E., Ansorge J. and Mueller S. 1998. Three-dimensional interface modelling with two-dimensional seismic data: the Alpine crust mantle boundary. *Geophysical Journal International* **135**, 264–278.
- Waldhauser F., Lippitsch R., Kissling E. and Ansorge J. 2002. High-resolution teleseismic tomography of upper-mantle structure using an a priori three-dimensional crustal model. *Geophysical Journal International* **150**, 403–414.
- Wapenaar K. 2004. Retrieving the elastodynamic Green's Function of an arbitrary inhomogeneous medium by cross-correlation. *Physical Review Letter* **93**, 254301.
- Wapenaar K. 2006. Nonreciprocal Green's function retrieval by cross correlation. *The Journal of the Acoustical Society of America* **120**, 1, EL7–E13.
- Weaver R.L. 2005. Information from seismic noise. *Science* **307**, 5715, 1568–1569.
- Weaver R.L. and Lobkis O.I. 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies. *Physical Review Letter* **87**, 13, 134301.
- Weaver R.L. and Lobkis O.I. 2003. Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal phonons. *The Journal of the Acoustical Society of America* **113**, 2611–2621.
- Weaver R.L. and Lobkis O.I. 2004. Diffuse fields in open systems and the emergence of the Green's function. *The Journal of the Acoustical Society of America* **116**, 5, 2731–2734.
- Weaver R.L. and Lobkis O.I. 2005a. Fluctuations in diffuse field-field correlations and the emergence of the Green's function in open systems. *The Journal of the Acoustical Society of America* **117**, 3432–3439.
- Weaver R.L. and Lobkis O.I. 2005b. The mean and variance of diffuse field correlations in finite bodies. *The Journal of the Acoustical Society of America* **118**, 3447–3456.
- Weaver R.L. and Lobkis O.I. 2006. Diffuse fields in ultrasonics and seismology. *Geophysics* **71**, SI5–SI9.
- Yang Y., Ritzwoller M.H., Levshin A.L. and Shapiro N.M. 2007. Ambient noise Rayleigh wave tomography across Europe. *Geophysical Journal International* **168**, 259–274.
- Yao H., van der Hilst R.D. and de Hoop M.V. 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps. *Geophysical Journal International* **166**, 732–744.
- Zollo A., D'Auria L., Matteis R.D., Herrero A., Virieux J. and Gasparini P. 2002. Bayesian estimation of 2-D P-velocity models from active seismic arrival time data: imaging of the shallow structure of Mt Vesuvius. *Geophysical Journal International* **151**, 566–582.

APPENDIX: DETAILED CALCULATION OF THE CROSS-CORRELATION FUNCTION (EQUATION 4)

We start from the definition of the cross-correlation function between two points A and B (equation 3) in which we express

the wavefield u using the Green's function G :

$$\begin{aligned} C(\tau, \vec{r}_A, \vec{r}_B) &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T u(t, \vec{r}_A) \overline{u(t + \tau, \vec{r}_B)} dt \\ &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T dt \int_0^\infty ds \int_X d\vec{r}_s G_a(s, \vec{r}_A, \vec{r}_s) f(t - s, \vec{r}_s) \\ &\quad \times \int_0^\infty ds' \int_X d\vec{r}_s' \overline{G_a(s', \vec{r}_B, \vec{r}_s') f(t + \tau - s', \vec{r}_s')} \end{aligned}$$

The large T limit can be replaced by an ensemble average, which gives the mathematical expectation denoted by \mathbb{E} . As f is a white noise, we have:

$$\begin{aligned} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(t - s, \vec{r}_s) f(t + \tau - s', \vec{r}_s') dt \\ = \mathbb{E}[f(t - s, \vec{r}_s) f(t + \tau - s', \vec{r}_s')] \\ = \sigma^2 \delta(\tau + s - s') \delta(\vec{r}_s - \vec{r}_s') \end{aligned}$$

where σ is the variance of the white noise. This property simplifies the previous equation, and we obtain:

$$C(\tau, \vec{r}_A, \vec{r}_B) = \sigma^2 \int_0^\infty ds \int_X d\vec{r}_s G_a(s, \vec{r}_A, \vec{r}_s) \overline{G_a(s + \tau, \vec{r}_B, \vec{r}_s)}$$

Using the expression of the Green's function (equation 2):

$$C(\tau, \vec{r}_A, \vec{r}_B) = \sigma^2 \int_0^\infty ds \int_X d\vec{r}_s Y(s) Y(s + \tau) e^{-as} e^{-a(s+\tau)}$$

$$\begin{aligned} &\times \left[\frac{\sin s \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_A, \vec{r}_s) \\ &\times \left[\frac{\sin(s+\tau) \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_B, \vec{r}_s) \end{aligned}$$

We use two properties of the integral kernel:

$$\overline{\|P\|}(x, y) = \overline{\|P\|}(x, y) = \|p\|(y, x)$$

$$\int_X \|P_1\|(x, z) \|P_2\|(z, y) dz = \|P_1 \cdot P_2\|(x, y)$$

to obtain a new formula for the cross-correlation function:

$$\begin{aligned} C(\tau, \vec{r}_A, \vec{r}_B) &= \sigma^2 \int_0^\infty ds Y(s + \tau) e^{-a(2s+\tau)} \\ &\times \left[\frac{\sin s \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \frac{\sin(s + \tau) \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right] (\vec{r}_A, \vec{r}_B) \end{aligned}$$

Using $\sin \alpha \sin \beta = 1/2 (\cos(\alpha - \beta) - \cos(\alpha + \beta))$ and computing the integral over ds , we obtain equation 4:

$$\begin{aligned} C(\tau, \vec{r}_A, \vec{r}_B) &= \frac{\sigma^2 e^{-a|\tau|}}{4a} \\ &\times \left[(-L)^{-1} \left(\cos \tau \sqrt{-L - a^2} + a \frac{\sin |\tau| \sqrt{-L - a^2}}{\sqrt{-L - a^2}} \right) \right] (\vec{r}_A, \vec{r}_B) \end{aligned}$$

Bibliographie générale

Aizawa, K., Yoshimura, R., Oshiman, N., Yamazaki, K., Uto, T., Ogawa, Y., Tank, S. B., Kanda, W., Sakanaka, S., Furukawa, Y., Hashimoto, T., Uyeshima, M., & Hurst, A. W., 2005. Hydrothermal system beneath Mt. Fuji volcano inferred from magnetotellurics and electric self-potential, *Earth & Planetary Science Letters*, **235**(1-2), 343–355.

Aki, K., 1957. Space and time spectra of stationary stochastic waves with special reference to microtremors, *Bulletin of the Earthquake Research Institute*, **35**, 415–456.

Aki, K., 1965. A note on the use of microseisms in determining the shallow structures of the Earth's crust, *Geophysics*, **30**(4), 665–666.

Aki, K., 1969. Analysis of the seismic coda of local earthquakes as scattered waves, *Journal of Geophysical Research*, **74**, 615–618.

Aki, K. & Chouet, B., 1975. Origin of coda waves: Source, attenuation, and scattering effects, *Journal of Geophysical Research*, **80**(9), 3322–3342.

Asten, M. W., 2006. On bias and noise in passive seismic data from finite circular array data processing using SPAC methods, *Geophysics*, **71**(6), V153–V162.

Asten, M. W. & Henstridge, J. D., 1984. Array estimators and use of microseisms for reconnaissance of sedimentary basins, *Geophysics*, **49**, 1828–1837.

Asten, M. W., Dhu, T., & Lam, N., 2004. Optimised array design for microtremor array studies applied to site classification; observations, results and future use, in *Proceedings of the 13th World Conference of Earthquake Engineering, Vancouver*.

Baggeroer, A. B., Kuperman, W. A., & Schmidt, H., 1988. Matched field processing: source localization in correlated noise as an optimum parameter estimation problem, *The Journal of the Acoustical Society of America*, **83**(2), 571–587.

Baggeroer, A. B., Kuperman, W. A., & Mikhalevsky, P. N., 1993. An overview of matched fields methods in ocean acoustics, *IEEE Journal of Ocean Engineering*, **18**(4), 401–424.

- Baig, A. & Campillo, M., 2008. Enhancing noise cross-correlations to identify higher-mode surface waves, *Geophysical Research Abstracts*, **10**, EGU2008–A–08541.
- Bedrosian, P. A., Unsworth, M. J., & Johnston, M. J., 2008. Hydrothermal circulation at Mount St. Helens determined by self-potential measurements, *Journal of Volcanology and Geothermal Research*, **160**, 137–146.
- Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophysical Journal International*, **169**(3), 1239–1260.
- Bettig, B., Bard, P.-Y., Scherbaum, F., Riepl, J., Cotton, F., Cornou, C., & Hatzfeld, D., 2001. Analysis of dense array noise measurements using the modified spatial auto-correlation method MSPAC – application to the Grenoble area, *Bollettino di Geofisica Teorica e Applicata*, **42**(3-4), 281–304.
- Bonafede, M. & Mazzanti, M., 1998. Modelling gravity variations consistent with ground deformation in the Campi Flegrei caldera (Italy), *Journal of Volcanology and Geothermal Research*, **81**, 137–157.
- Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E., 2008a. Postseismic relaxation along the San Andreas fault in the Parkfield area investigated with continuous seismological observations, *Science*, p. submitted.
- Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, Z., Duputel, Z., Coutant, O., & Nercessian, A., 2008b. Towards forecasting volcanic eruptions using seismic noise, *Nature Geoscience*, **1**(1), 126–130.
- Bruno, P. P. G., Ricciardi, G. P., Petrillo, Z., Di Fiore, V., Troiano, A., & Chiodini, G., 2007. Geophysical and hydrogeological experiments from a shallow hydrothermal system at Solfatara Volcano, Campi Flegrei, Italy: response to caldera unrest, *Journal of Geophysical Research (Solid Earth)*, **112**(B6), 6201.
- Burg, J. P., 1964. Three-dimensional filtering with array of seismometers, *Geophysics*, **29**, 693–713.
- Campillo, M., 2006. Phase and correlation in ‘random’ seismic fields and the reconstruction of the Green function, *Pure and Applied Geophysics*, **163**(2-3), 475–502.
- Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**, 547–549.
- Campillo, M., Margerin, L., & Shapiro, N. M., 1999. *Seismic wave diffusion in the Earth lithosphere*, NATO ASI Series, Kluwer.
- Campman, X. H., van Wijk, K., Scales, J. A., & Herman, G. C., 2005. Imaging and suppressing near-receivers scattered surface waves, *Geophysics*, **70**(2), V21–V29.

- Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proceedings of the IEEE*, **57**(8), 1408–1418.
- Cassereau, D. & Fink, M., 1992. Time-reversal of ultrasonic fields – Part III: Theory of the closed time-reversal cavity, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **39**(5), 579–592.
- Chávez-García, F. J., Rodríguez, M., & Stephenson, W. R., 2005. An alternative approach to the SPAC analysis of microtremors: exploiting stationarity of noise, *Bulletin of the Seismological Society of America*, **95**(1), 277–293.
- Cho, I., Tada, T., & Shinozaki, Y., 2004. A new method to determine phase velocities of Rayleigh waves from microseisms, *Geophysics*, **69**(6), 1535–1551.
- Cho, I., Tada, T., & Shinozaki, Y., 2006. A generic formulation for microtremor exploration methods using three-components records from a circular array, *Geophysical Journal International*, **165**, 236–258.
- Claerbout, J. F., 1968. Synthesis of a layered medium from its acoustics transmission response, *Geophysics*, **33**, 264–269.
- Colin de Verdière, Y., 2006a. Mathematical models for passive imaging I: general background.
- Colin de Verdière, Y., 2006b. Mathematical models for passive imaging II: effective hamiltonians associated to surface waves.
- Colin de Verdière, Y., 2008. Semi-classical analysis and passive imaging, Submitted to Nonlinearity.
- Cornou, C., Ohrnberger, M., Boore, D. M., Kudo, K., & Bard, P.-Y., 2006. Using ambient noise array techniques for site characterizations: results from an international benchmark, in *Proceedings of the 3rd International Symposium on the Effects of Surface Geology on Seismic Motion, Grenoble, 29 August – 01 September 2006*, LCPC Editions.
- Cowan, M. L., Page, J. H., & Weitz, D. A., 2000. Velocity fluctuations in fluidized suspensions probed by ultrasonic correlation spectroscopy, *Physical Review Letters*, **85**(2), 453–456.
- Cox, H., Zeskind, R. M., & Owen, M. M., 1987. Robust adaptive beamforming, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-35**(10), 1365–1376.
- de Rosny, J. & Roux, P., 2001. Multiple scattering in a reflecting cavity: Application to fish counting in a tank, *The Journal of the Acoustical Society of America*, **109**(6), 2587–2597.
- de Rosny, J. & Roux, P., 2004. Reply to the comment on "Multiple scattering in a reflecting cavity: Application to fish scattering" [J. Acoust. Soc. Am. 113, 2978-2979 (2003)] (1), *The Journal of the Acoustical Society of America*, **115**(1), 31–34.
- Debever, C. & Kuperman, W. A., 2007. Robust matched-field processing using a coherent broadband white noise constraint processor, *The Journal of the Acoustical Society of America*, **122**(4), 1979–1986.

- Derode, A., Tourin, A., & Fink, M., 2001. Random multiple scattering of ultrasound. I. Coherent and ballistic waves, *Physical Review E*, **64**(3), 036605.
- Derode, A., Larose, E., Campillo, M., & Fink, M., 2003a. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Applied Physics Letters*, **83**(15), 3054–3056.
- Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M., & Fink, M., 2003b. Recovering the Green's function from field-field correlations in an open scattering medium (L), *The Journal of the Acoustical Society of America*, **113**(6), 2973–2976.
- Draganov, D., Wapenaar, K., Mulder, W., Singer, J., & Verdel, A., 2007. Retrieval of reflections from seismic background-noise measurements, *Geophysical Research Letters*, **34**, L04305.
- Duputel, Z., Ferrazzini, V., Brenguier, F., Shapiro, N. M., Campillo, M., & Nercessian, A., 2008. Real time monitoring of relative velocity changes using ambient seismic noise on the Piton de la Fournaise volcano (La Reunion) from January 2006 to June 2007, *Journal of Volcanology and Geothermal Research*, p. submitted.
- Duvall, T. L., Jefferies, S. M., Harvey, J. W., & Pomerantz, M. A., 1993. Time distance helioseismology, *Nature*, **362**, 430–432.
- Fink, M., 1992. Time reversal of ultrasonic fields – Part I: Basic principles, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **39**(5), 555–566.
- Fink, M., Cassereau, D., Derode, A., Prada, C., Roux, P., Tanter, M., Thomas, J.-L., & Wu, F., 2000. Time-reversed acoustics, *Reports on Progress in Physics*, **63**, 1933–1995.
- Gaffet, S., Larroque, C., Deschamps, A., & Tressols, F., 1998. A dense array experiment for the observation of waveform perturbations, *Soil Dynamics and Earthquake Engineering*, **17**, 475–484.
- Gilles, P. M., Duvall, T. L., Scherrer, P. H., & Bogart, R. S., 1997. A subsurface flow of material from the Sun's equator to its poles, *Nature*, **390**, 52–54.
- Godin, O. A., 2007. Emergence of the acoustic Green's function from thermal noise, *The Journal of the Acoustical Society of America*, **121**(2), EL96–EL102.
- Gouédard, P., Roux, P., & Campillo, M., 2006. Dispersion curves measurement using noise cross-correlation technique, in *Proceedings of the 3rd International Symposium on the Effects of Surface Geology on Seismic Motion, Grenoble, 29 August – 01 September 2006*, LCPC Editions, paper N06.
- Gouédard, P., Cornou, C., & Roux, P., 2008a. Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique, *Geophysical Journal International*, **172**(3), 971–981.
- Gouédard, P., Roux, P., & Campillo, M., 2008b. Small-scale seismic inversion using surface waves extracted from noise cross correlation, *The Journal of the Acoustical Society of America, Express Letters*, **123**(3), EL26–EL31.

- Gouédard, P., Roux, P., Campillo, M., & Verdel, A., 2008c. Convergence of the two-point correlation function toward the Green's function in the context of a seismic prospecting dataset, *Geophysics*, **73**(6), V47–V53.
- Gouédard, P., Stehly, L., Brenguier, F., Campillo, M., Colin de Verdière, Y., Larose, E., Margerin, L., Roux, P., Sánchez-Sesma, F. J., Shapiro, N. M., & Weaver, R. L., 2008d. Cross-correlation of random fields: mathematical approach and applications, *Geophysical Prospecting*, **56**(3), 375–393.
- Hennino, R., Tégourès, N., Shapiro, N. M., Margerin, L., Campillo, M., van Tiggelen, B. A., & Weaver, R. L., 2001. Observation of equipartition of seismic waves, *Physical Review Letters*, **86**(15), 3447–3450.
- Henstridge, J. D., 1979. A signal processing method for circular arrays, *Geophysics*, **44**, 179–184.
- Herman, G. C. & Perkins, C., 2006. Predictive removal of scattered noise, *Geophysics*, **71**(2), V41–V49.
- Herrmann, R. B., 1987. Computer programs in seismology, Department of Earth and Atmospheric Sciences, Saint Louis University, Saint Louis, available for free at <http://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html>.
- Hisada, Y., 1994. An efficient method for computing Green's functions for a layered halfspace with sources and receivers at close depths, *Bulletin of the Seismological Society of America*, **84**, 1456–1472.
- Hisada, Y., 1995. An efficient method for computing Green's functions for a layered halfspace with sources and receivers at close depths (part 2), *Bulletin of the Seismological Society of America*, **85**(4), 1080–1093.
- Horike, M., 1985. Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas, *Journal of Physics of the Earth*, **33**(2), 59–96.
- Iuliano, T., Mauriello, P., & Patella, D., 2002. Looking inside Mount Vesuvius by potential fields integrated probability tomographies, *Journal of Volcanology and Geothermal Research*, **113**(3), 363–378.
- Kanai, K. & Tanaka, T., 1954. Measurement of microtremor, *Bulletin of the Earthquake Research Institute*, **32**, 199–209.
- Keam, R. F., 1988. *Tarawera: the volcanic eruption of 10 June 1986*, R. F. Keam.
- Kedar, S., Sturtevant, B., & Kanamori, H., 1996. The origin of harmonic tremor at Old Faithful geyser, *Nature*, **379**, 708–711.
- Kedar, S., Kanamori, H., & Sturtevant, B., 1998. Bubble collapse as the source of tremor at Old Faithful geyser, *Journal of Geophysical Research*, **103**, 24283–24300.
- Kubo, R., 1966. The fluctuation-dissipation theorem, *Reports on Progress in Physics*, **29**(1), 255–284.

- Kuperman, W. A. & Turek, G., 1997. Matched field acoustics, *Mechanical Systems and Signal Processing*, **11**(1), 141–148.
- Lacoss, R. T., Kelly, E. J., & Toksöz, M. N., 1969. Estimation of seismic noise structure using arrays, *Geophysics*, **34**(1), 21–38.
- Larose, E., Derode, A., Campillo, M., & Fink, M., 2004. Imaging from one-bit correlation of wide-band diffuse wavefield, *Journal of Applied Physics*, **95**(12), 8393–8399.
- Larose, E., Khan, A., Nakamura, Y., & Campillo, M., 2005. Lunar subsurface investigated from correlation of seismic noise, *Geophysical Research Letters*, **32**, L16201.
- Larose, E., Lobkis, O. I., & Weaver, R. L., 2006a. Passive correlation imaging of a buried scatterer, *The Journal of the Acoustical Society of America*, **119**(6), 3549–3552.
- Larose, E., Margerin, L., Derode, A., van Tiggelen, B. A., Campillo, M., Shapiro, N. M., Paul, A., Stehly, L., & Tanter, M., 2006b. Correlation of random wavefields: an interdisciplinary review, *Geophysics*, **71**(4), SI11–SI21.
- Larose, E., Roux, P., & Campillo, M., 2007. Reconstruction of Rayleigh-Lamb dispersion spectrum based on noise obtained from an air-jet forcing, *The Journal of the Acoustical Society of America*, **122**(6), 3437–3444.
- Legaz, A., Revil, A., Roux, P., Vandemeulebrouck, J., Gouédard, P., Hurst, T., Bolève, A., Legaz, A., Revil, A., Roux, P., Vandemeulebrouck, J., Gouédard, P., Hurst, T., & Bolève, A., 2009. Self-potential and passive seismic monitoring of hydrothermal activity: A case study at Iodine Pool, Waimangu geothermal valley, New Zealand, *Journal of Volcanology and Geothermal Research*, **170**(1-2), 11–18.
- Lin, F., Ritzwoller, M. H., Townend, J., Bannister, S., & Savage, M., 2007. Ambient noise Rayleigh wave tomography of New Zealand, *Geophysical Journal International*, **170**(2), 649–666.
- Lobkis, O. I. & Weaver, R. L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *The Journal of the Acoustical Society of America*, **110**(6), 3011–3017.
- Louie, J., 2001. Faster, better: shear-wave velocity to 100 meters depth from refraction microtremor array, *Bulletin of the Seismological Society of America*, **91**, 347–364.
- Luo, Y., Xia, J., Liu, J., Xu, Y., & Liu, Q., 2008. Generation of a pseudo-2d shear-wave velocity section by inversion of a series of 1d dispersion curves, *Journal of Applied Geophysics*, **64**(3-4), 1115–1124.
- Margerin, L., Campillo, M., Shapiro, N. M., & van Tiggelen, B. A., 1999. Residence time of diffuse waves in the crust as a physical interpretation of coda Q : application to seismograms recorded in Mexico, *Geophysical Journal International*, **138**(2), 343–352.
- Margerin, L., Campillo, M., & van Tiggelen, B. A., 2000. Monte Carlo simulation of multiple scattering of elastic waves, *Journal of Geophysical Research*, **105**(B4), 7873–7892.

- Miyakoshi, K., 1996. A range of wavelengths possible to estimate phase velocities of surface waves in microtremors, in *Proceedings of the 94th SEGJ Conference, Society of Exploration Geophysicists, Japan*, pp. 178–182, (in Japanese).
- Moczo, P. & Kristek, J., 2002. Fd code to generate noise synthetics, in *Sesame report*, no. D02.09, available on <http://sesame-fp5.obs.ujf-grenoble.fr/>.
- Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface using microtremors on the ground surface, *Quarterly Reports of Railway Technological Institute of Tokyo*, **30**, 25–33.
- Nogoshi, M. & Igarashi, T., 1971. On the amplitude characteristics of microtremor (part 2), *Journal of the Seismological Society of Japan*, **24**, 26–40, (in Japanese with English abstract).
- Ohori, M., Nobata, A., & Wakamatsu, K., 2002. A comparison of ESAC and FK methods of estimating phase velocity using arbitrary shaped microtremor array, *Bulletin of the Seismological Society of America*, **92**(6), 2323–2332.
- Ohrnberger, M., 2005. Report on the FK/SPAC capabilities and limitations, in *SESAME Deliverable*, no. D19.06, pp. 43, <http://sesame-fp5.obs.ujf-grenoble.fr/Delivrables/Del-D19-Wp06.pdf>.
- Okada, H., 2003. The microseismic survey method, in *Geophysical Monographs*, vol. 12, Society of Exploration Geophysicists of Japan, translated by Koya Suto, Society of Exploration Geophysicists, Tulsa.
- Okada, H., 2006. Theory of efficient array observations of microtremors with special reference to the SPAC method, *Exploration Geophysics*, **37**, 73–85.
- Park, C. B., Miller, R. D., & Xia, J., 1999. Multichannel analysis of surface waves, *Geophysics*, **64**(3), 800–808.
- Patella, D., 1997. Self-potential global tomography including topographic effects, *Geophysical Prospecting*, **45**, 843–863.
- Paul, A. & Campillo, M., 2001. Extracting the Green function between two stations from coda waves, in *AGU Fall Meeting Abstracts*, p. D610.
- Paul, A., Campillo, M., Margerin, L., Larose, E., & Derode, A., 2005. Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves, *Journal of Geophysical Research (Solid Earth)*, **110**, B08302.
- Pedersen, H. A., Krüger, F., & the SVEKALAPKO Seismic Tomography Working Group, 2007. Influence of the seismic noise characteristics on noise correlations in the Baltic Shield, *Geophysical Journal International*, **168**(1), 197–210.
- Poupinet, G., Ellsworth, W. L., & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *Journal of Geophysical Research*, **89**(B7), 5719–5731.
- Ramm, A. G., 1986. *Scattering by obstacles*, Kluwer Academic Publisher.

- Reed, M. & Simon, B., 1978. *Methods of modern mathematical physics I, II, III, IV*, Academic Press.
- Revil, A., Ehouarne, L., & Thyreault, E., 2001. Tomography of self-potential anomalies of electrochemical nature, *Geophysical Research Letters*, **28**(23), 4363–4366.
- Revil, A., Finizola, A., Piscitelli, S., Rizzo, E., Ricci, T., Crespy, A., Angeletti, B., Balasco, M., Barde Cabusson, S., Bennati, L., Bolève, A., Byrdina, S., Carzaniga, N., Di Gangi, F., Morin, J., Perrone, A., Rossi, M., Roulleau, E., & Suski, B., 2008. Inner structure of La Fossa di Vulcano (Vulcano Island, southern Tyrrhenian Sea, Italy) revealed by high-resolution electric resistivity tomography coupled with self-potential, temperature, and CO₂ diffuse degassing measurements, *Journal of Geophysical Research*, **113**, B07207.
- Rickett, J. & Claerbout, J. F., 1999. Acoustic daylight imaging via spectral factorization; helioseismology and reservoir monitoring, *The Leading Edge*, **18**(8), 957–960.
- Roberts, P. M., Phillips, W. S., & Fehler, M. C., 1992. Development of the active doublet method for measuring small velocity and attenuation changes in solids, *The Journal of the Acoustical Society of America*, **91**(6), 3291–3302.
- Roux, P. & Kuperman, W. A., 2004. Extracting coherent wavefronts from acoustic ambient noise in the ocean, *The Journal of the Acoustical Society of America*, **116**(4), 1995–2003.
- Roux, P., Sabra, K. G., Gerstoft, P., & Kuperman, W. A., 2005a. P-waves from cross-correlation of seismic noise, *Geophysical Research Letters*, **32**, L19303.
- Roux, P., Sabra, K. G., Kuperman, W. A., & Roux, A., 2005b. Ambient noise cross correlation in free space: theoretical approach, *The Journal of the Acoustical Society of America*, **117**(1), 79–84.
- Ryzhik, L., Papanicolaou, G., & Keller, J. B., 1996. Transport equations for elastic and other waves in random media, *Wave motion*, **24**, 327–370.
- Sabra, K. G., Gerstoft, P., Roux, P., & Kuperman, W. A., 2005a. Surface wave tomography from microseisms in southern California, *Geophysical Research Letters*, **32**, L14311.
- Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C., 2005b. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophysical Research Letters*, **32**, L03310.
- Sabra, K. G., Roux, P., & Kuperman, W. A., 2005c. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide, *The Journal of the Acoustical Society of America*, **117**(1), 164–174.
- Sabra, K. G., Roux, P., & Kuperman, W. A., 2005d. Emergence rate of the time-domain Green's function from the ambient noise cross-correlation, *The Journal of the Acoustical Society of America*, **118**(6), 3524–3531.
- Sabra, K. G., Srivastava, A., di Scalea, F. L., Bartoli, I., Rizzo, P., & Conti, S., 2008. Structural health monitoring by extraction of coherent guided waves from diffuse fields, *The Journal of the Acoustical Society of America*, **123**(1), EL8–EL13.

- Sánchez-Sesma, F. J. & Campillo, M., 2006. Retrieval of the Green function from cross correlation: the canonical elastic problem, *Bulletin of the Seismological Society of America*, **96**(3), 1182–1191.
- Sánchez-Sesma, F. J., Pérez-Ruiz, J. A., Campillo, M., & Luzón, F., 2006. Elastodynamic 2-D Green function retrieval from cross-correlation: canonical inclusion problem, *Geophysical Research Letters*, **33**, L13305.
- Sánchez-Sesma, F. J., Pérez-Ruiz, J. A., Luzón, F., Campillo, M., & Rodríguez-Castellanos, A., 2008. Diffuse fields in dynamic elasticity, *Wave Motion*, **45**(5), 641–654.
- Satoh, T., Kawase, H., Iwata, T., Higashi, S., Sato, T., Irikura, K., & Huang, H.-C., 2001. A-wave velocity structure of the Taichung basin, Taiwan, estimated from array and single-station records of microtremors, *Bulletin of the Seismological Society of America*, **91**(5), 1267–1282.
- Schuster, G. T., Yu, J., Sheng, J., & Rickett, J., 2004. Interferometric/daylight seismic imaging, *Geophysical Journal International*, **157**(2), 838–852.
- Scott, B., 1992. Characteristics of cyclic activity in Frying Pan and Inferno crater lakes, Waimangu, in *14th New Zealand Geothermal Workshop*, Geothermal Institute, University of Auckland, New Zealand, pp. 253–258.
- Scott, B., 1994. Cyclic activity in the crater lakes of Waimangu hydrothermal system, New-Zealand, *Geothermics*, **23**, 555–572.
- Shapiro, N. M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophysical Research Letters*, **31**, L07614.
- Shapiro, N. M., Campillo, M., Margerin, L., Singh, S. K., Kostoglodov, V., & Pacheco, J., 2000. The energy partitioning and the diffusive character of the seismic coda, *Bulletin of the Seismological Society of America*, **90**(3), 655–665.
- Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.
- Shuli, W. & Zhuo, L., 2003. Water hammer phenomena in gas-water two-phase bubbly flow through a 90-degree bend tube, *Journal of Fluids Engineering*, **125**, 736–737.
- Snieder, R., 2002. Coda wave interferometry and the equilibration of energy in elastic media, *Physical Review E*, **66**, 046615.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Physical Review E*, **69**(4), 046610.
- Snieder, R., 2006. The theory of coda wave interferometry, *Pure and Applied Geophysics*, **163**(2), 455–473.
- Stanton, T., 1978. *Seismic noise characteristics of Inferno Crater Lake, Waimangu*, Msc thesis, Auckland University, Auckland.

- Stehly, L., Campillo, M., & Shapiro, N. M., 2006. A study of the seismic noise from its long range correlation properties, *Journal of Geophysical Research*, **111**, B10306.
- Stehly, L., Campillo, M., & Shapiro, N. M., 2007. Travel time measurements from noise correlation: stability and detection of instrumental time-shifts, *Geophysical Journal International*, **171**(1), 223–230.
- Tarantola, A., 1987. *Inverse problem theory: methods for data fitting and model parameter estimation*, Amsterdam.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Reviews of Geophysics and Space Physics*, **20**, 219–232.
- Tokimatsu, K., 1995. Geotechnical site characterization using surface waves, in *Proceedings of the 1st International Conference of Earthquake Geotechnical Engineering*, pp. 1333–1368, Balkema.
- Tosha, T., Ishido, T., Nishi, Y., Matsushima, N., & Scott, B., 1996. Seismological and self-potential surveys at Inferno crater lake, Waimangu geothermal field, in *18th New Zealand Geothermal Workshop*, Geothermal Institute, University of Auckland, New Zealand, pp. 249–254.
- Tourin, A., Derode, A., & Fink, M., 2001. Sensitivity to perturbations of a time-reversed acoustic wave in a multiple scattering medium, *Physical Review Letters*, **87**(27), 274301.
- van Tiggelen, B. A., 2003. Green function retrieval and time reversal in a disordered world, *Physical Review Letters*, **91**(24), 243904.
- Vandemeulebrouck, J., Stemmelen, D., Hurst, T., & Grangeon, J., 2005. Analogue modeling of instabilities in crater lake hydrothermal systems, *Journal of Geophysical Research (Solid Earth)*, **110**(B2), B02212.
- Vandemeulebrouck, J., Gouédard, P., Legaz, A., Roux, P., Revil, A., Hurst, A. W., Jardani, A., & Bolève, A., 2009. Application of seismic noise and self-potential localization techniques to a buried hydrothermal vent (Waimangu Old Geyser site, New-Zealand), *Geophysical Journal International*, *Fast Track*, **submitted**.
- Wapenaar, K., 2003. Synthesis of an inhomogeneous medium from its acoustic transmission response, *Geophysics*, **68**(5), 1756–1759.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's Function of an arbitrary inhomogeneous medium by cross-correlation, *Physical Review Letters*, **93**(25), 254301.
- Wapenaar, K., 2006. Nonreciprocal Green's function retrieval by cross correlation, *The Journal of the Acoustical Society of America*, **120**(1), EL7–EL13.
- Wapenaar, K., Fokkema, J., & Snieder, R., 2005. Retrieving the Green's function by cross-correlation: a comparison of approaches (I), *The Journal of the Acoustical Society of America*, **118**(5), 2783–2786.
- Wathelet, M., 2008. An improved neighborhood algorithm: parameter conditions and dynamic scaling, *Geophysical Research Letters*, **35**, L09301.

- Wathelet, M., Jongmans, D., & Ohrnberger, M., 2004. Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surface Geophysics*, **2**, 211–221.
- Wathelet, M., Jongmans, D., Ohrnberger, M., & Bonnefoy-Claudet, S., 2008. Array performances for ambient vibrations on a shallow structure and consequences over V_s inversion, *Journal of Seismology*, **12**(1), 1–19.
- Weaver, R. L., 2005. Information from seismic noise, *Science*, **307**(5715), 1568–1569.
- Weaver, R. L. & Lobkis, O. I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Physical Review Letters*, **87**(13), 134301.
- Weaver, R. L. & Lobkis, O. I., 2003. Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal phonons, *The Journal of the Acoustical Society of America*, **113**(5), 2611–2621.
- Weaver, R. L. & Lobkis, O. I., 2004. Diffuse fields in open systems and the emergence of the Green's function, *The Journal of the Acoustical Society of America*, **116**(5), 2731–2734.
- Weaver, R. L. & Lobkis, O. I., 2005a. Fluctuations in diffuse field-field correlations and the emergence of the Green's function in open systems, *The Journal of the Acoustical Society of America*, **117**(6), 3432–3439.
- Weaver, R. L. & Lobkis, O. I., 2005b. The mean and variance of diffuse field correlations in finite bodies, *The Journal of the Acoustical Society of America*, **118**(6), 3447–3456.
- Woods, J. W. & Lintz, P. L., 1973. Plane waves at small arrays, *Geophysics*, **38**, 1023–1041.
- Yao, H., van der Hilst, R. D., & de Hoop, M. V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps, *Geophysical Journal International*, **166**(2), 732–744.